

Disappointing end for the SSC

The decision of the US Congress to kill the Superconducting Super Collider does not and should not spell the death of particle physics, but rather the beginning of a strictly international programme.

THE traumatic end of the Superconducting Super Collider (SSC) last week is a comprehensive tragedy. There are no winners, only losers. The decision (see page 773) by the US House of Representatives to reaffirm its earlier opposition to the Department of Energy's budget request is not, as some may claim, a "victory for small science", whatever that may be; the funds now saved will not be redistributed among small groups of dedicated people working in ordinary laboratories, but will instead tumble into the black hole of the US federal deficit and will not be seen again.

Nor is the decision much comfort for CERN, the European particle physics laboratory at Geneva. Its member governments may now sense that the US decision eases the pressure for a quick decision on the Large Hadron Collider (LHC), the plan to build a superconducting proton-antiproton collider in the same tunnel as the existing electron-positron collider (LEP). But the real losers, who deserve the whole research community's compassion, are those who, for more than ten years, have worked to make a reality of the SSC. The innumerable design panels have taken years of many able people's work. The 1,000 or so who hitched their careers to the SSC by moving to Texas are especially unfortunate; three months' pay will not be much compensation.

What went wrong? It will be tempting, but mistaken, to blame the US Congress for shortsightedness. Although the Congress is known to have the federal deficit on its mind, its conversion to the administration's view that research spending should be harnessed strictly to industrial competitiveness is a little more surprising. And in any case, is not the Congress traditionally exempt from the requirements of consistency when translating principles into practice? And why should it now kill the SSC, which has definable if esoteric objectives, when it has agreed to keep alive for another year the space station Freedom, which has none? Sadly, none of these questions carries weight. The truth is that the SSC proposal has embodied a fatal flaw from the outset.

Flaw

The flaw is not technical but structural and diplomatic. There can be no substantial quarrel with the goals of particle physics, to understand what matter is made of and, more important, why its ingredients are what they are. Experimental particle physics has its roots in the United States in the heroic 1930s. Since the Second World War, US accelerators have blazed the trail of discovery by arranging for ever-more energetic collisions between particles. In Europe, it was

soon (in the 1960s) recognized that national programmes could not keep up, but the European consortium CERN has become an increasingly important source of discovery. (Accelerators in Germany, Japan and the former Soviet Union have also made important contributions.) But costs have mounted steadily all the while. And even in the United States, \$11 billion is now a lot of money. SSC is dead because the US Congress, the arm of government that decides what the United States can afford to spend, has now followed other governments in deciding that particle physics is too costly to be a national pursuit.

Blame

The leaders of the US particle physics community must blame themselves for not appreciating that such a point would eventually be reached. To be fair, for at least a decade they have acknowledged that particle accelerators would eventually have to be built and operated internationally. But, echoing St Augustine's prayer on chastity ("Lord make me chaste, but not yet!"), have successfully argued that they should first build one more machine. Not only the US deficit but also the disappearance of the Cold War's false criteria of leadership have now caught them out. The often unseemly efforts to give SSC the appearance of an international project (especially in linking the demand for a Japanese contribution to bilateral trade talks) were both demeaning and unconvincing. Why build SSC when LHC could either demonstrate that the top quark and the Higgs boson exist or suggest, more accurately than SSC's designers could have known ten years ago, how they might be found?

So what will happen now to the particle physics community in the United States? Everybody's interest is that should remain in being and not become demoralized. Luckily, there are three laboratories (Fermilab and at Stanford and Cornell) at which experimental people can still do important work. But the United States should also pursue the option of a deal with CERN. The ideal would be full membership, but that will not be quickly brought about. (CERN is constituted by an international treaty, adherence to which would require the approval of the US Senate.) But that is the best way in which the US community can help to decide what kind of particle accelerator follows LHC, perhaps a pair of linear accelerators colliding electrons and positrons.

Meanwhile, the SSC calamity raises questions of how the United States should continue to arrange for the support of science. Particle physics has been for 50 years dependent on



the US Atomic Energy Commission and then the Department of Energy. Because departments of government fight for what they have, particle physics has been supported generously, and often in the face of resentment by those who must struggle with the indignities of peer-review. But the downside of that privilege, now apparent, is that there has been no shield between particle physics and the Congress. When notoriously fickle representatives could be counted on to regard particle physics as an element in the struggle with the Soviet Union, that did not matter. Now, for all the extra trouble that would be involved, the National Science Foundation might be a better sponsor for the long run. —

Is virtual reality real?

Virtual reality may be more than an extension of the cinema, but quite what is not yet clear.

WHAT is virtual reality if it is not simply a verbal joke or a pursuit for teenagers in games arcades? Plainly, there is more to be said of it than can be learned from young people wearing strange helmets about the sensory organs carried externally on their heads. Two weeks ago, the US Institution of Electrical and Electronic Engineers (IEEE) held a meeting in New York intended to help endow virtual reality with professional clothes. Separately, it has formed technical and standards committees to give shape to this emerging field, somewhere between psychology and computer science. But just what virtual reality may be is still not crystal clear.

Whatever it is, virtual reality has its roots in the age-old craft of manufacturing illusion, of which it might be held that the classical high-point is the cinema; a person spellbound for an hour and a few minutes by *High Noon* is indeed immersed in a reality that never was, a railway town in the Western desert of the United States a century ago. It hardly matters that the film is not in colour; indeed, that may help a person to distinguish the real reality from the false. But now, of course, the technical means of representing reality are far superior and are no longer confined to the cinema screen, but may be directly linked to the human brain's own input-output devices, slow (by mechanical standards) though they may be. But that is only a way of going to the cinema by putting a helmet on your head.

True virtual relationists have more powerful ambitions. Beyond immersion in a manufactured reality comes the ability to move about within it (or to change perspective) and, then, to interact with it. Kicha Ganapathy from AT&T Bell Laboratories said in New York that interaction is essential, realism less so. Although these criteria may be satisfied in playing chess against an opponent no more substantial than a computer program, more than that is intended. The manufactured reality must be a representation of a manufactured world, complexity and all.

Where will all this lead? Arcade games are already on the way. Psychologists will no doubt learn much of people's behaviour in variously manufactured worlds. Teaching people skills, from driving cars and piloting aircraft to speaking foreign languages conversationally, is within the sights of

the helmet-builders. But is it possible that the explanation of the interest of the major telecommunications companies in virtual reality has something to do with the prospect that when US Vice President Al Gore has provided the United States with his promised electronic superhighway, the need (and opportunity) will be to find ways of using that vast communications capacity? Arno Penzias, research director at AT&T, almost admitted as much in New York. □

Step back in Ulster

After another bomb, long-term institution building should be the goal in Northern Ireland.

WHAT is to be done about Northern Ireland? Since the present wave of troubles began in 1969, that question has provoked nothing but despair from the governments of the United Kingdom and the Irish Republic, not to mention those who live in the province and the many battalions of British troops drafted there in a hopeless attempt to keep the peace. But for the past few weeks, there have been repeated suggestions from official and unofficial sources that constructive talks were under way. What, for Northern Ireland, more natural than that the latest act of violence should have been last Saturday? The explosion of a bomb in a fishmonger's shop in the Protestant district of Belfast killed ten people. Talk of truce, even armistice, has fallen off while Roman Catholics brace themselves for retaliation.

This is not the first time that hopes have been dashed in Northern Ireland. Nor is it likely to be the last; the roots of the conflict are too ancient for that. But the present dynamics are all too clear. The Irish Republican Army (IRA), illegal both in Northern Ireland and in the republic, wants to drive the British Army out of Northern Ireland; its more distant goal of a united Ireland is not, as it would be delivered, likely to be welcome to the government in Dublin. The activities of the IRA provoked the existence of retaliatory (and also illegal) armed bands, now free-standing, which are capable of equally ghastly murders. It is not known to what extent the sustenance of the opposing groups by the communities they claim to represent is voluntary or coercive, but each has some of the attributes of a protection racket.

It is appalling that such circumstances should persist in an otherwise cultivated society. The present British and Irish governments are cooperating fruitfully on immediate and even some distant problems. But in case immediate solutions are beyond reach, they need a recipe for the more distant future. Because community separation is reinforced by sectarian education, should that not be abolished in Northern Ireland — even if that means abolishing sectarian education in the rest of the United Kingdom? And why not build on an imaginative scheme by the University of Ulster to build a new campus in Belfast in part of a no-man's-land separating the two communities by integrating all of Ireland's universities into a single system? Then at least there would emerge a nucleus of graduates persuaded that there is more to Irish life than its divisions. And there are many other institutions that could be denationalized in this way. □

SSC decision ends post-war era of science-government partnership

Washington. A physics community stunned by Congress's final decision to kill the Superconducting Super Collider (SSC) has been waking up to the realization that, both for its own discipline and for science in general, nothing will ever be quite the same again.

"This [decision] breaks the very successful partnership between science and government which has extended since the war, starting from the Manhattan Project", says Stan Wojcicki of Stanford University, chairman of the government's High Energy Physics Advisory Panel (HEPAP). "That partnership involved a level of mutual confidence which has now broken down."

Last Thursday, 21 October, Hazel O'Leary, the secretary of energy, issued a shock order terminating the \$11 billion Texan project — which is one-fifth complete — following a decision by a joint committee of the Senate and House of Representatives to pull the plug.

The decision was taken after powerful supporters in the Senate, led by Bennett Johnston (Democrat, Louisiana), conceded to overwhelming opposition in the House, where critics had claimed that the project, however worthy, was unaffordable.

Some SSC supporters blamed the Clinton administration's half-hearted support for the project's collapse. Although the administration has supported the project in public, it is said to be privately relieved to watch it die, as it tries to shift spending on science from basic to applied research (see *Nature* 365, 195: 1993).

As the Internet communications network surged with tens of thousands of messages between anxious scientists whose work relates to the project, leading US high energy physicists were regrouping to ensure the survival of their discipline by going international — probably through greater participation in the proposed Large Hadron Collider (LHC) at CERN, the European Laboratory for Particle Physics.

"The only good thing to come out of this will be more international collaboration", says Sidney Drell, deputy director of the Stanford Linear Accelerator Center (SLAC) in California. Drell urges physicists to work for fully-fledged US and Japanese involvement to convert CERN from a European centre to an international one.

"It's important that we move forward while there is still some momentum on this frontier of research, and the LHC is the one machine that we know how to build", says Drell. "The most horrible thing would be if we just folded up the tent and went home."

George Brown (Democrat, California), who is chairman of the House Science Com-

mittee, describes the decision as a "tragedy", and says that his committee will now seek to engage the White House's Office of Science and Technology Policy and O'Leary's Energy Department in a "serious

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

exploration" of future US involvement in fundamental physics.

"The only course I can see now is one of international collaboration, possibly on the CERN upgrade", says Brown. He adds that a formal affiliation to CERN would "not be a big step for us". CERN's charter currently

CERN sees move as mixed blessing

Munich. When Carlo Rubbia, head of the European Laboratory for Particle Physics (CERN) in Geneva, formally put forward his original plan to build the Large Hadron Collider (LHC) two years ago, the organization's 19 member states unanimously agreed that it was the right way forward for European high energy physics.

Washington's decision last week to stop construction of the equivalent US machine, the Superconducting Super Collider (SSC), may have shaken this confidence. But CERN physicists remain confident that the LHC will indeed be approved next spring, despite rumoured increases in cost.

The Rubbia plan anticipates that the LHC will absorb virtually all CERN's resources during its development, scheduled for completion in 2001. The total cost of the basic collider of between SFr2 and SFr3 billion (US\$1.4–\$2.1 billion) would come out of CERN's normal subscriptions; its two detectors, costing another SFr700 million, would be funded from other sources, in particular national science agencies.

According to supporters of the project, the steep bill is justified by the clear scientific aims of the project, including the detection of the predicted fundamental particle known as the Higgs boson.

Work on the design of the LHC and its detectors has been under way for some time.

excludes non-European countries from full membership, although 500 American and 110 Japanese physicists work at the facility, on the Swiss-French border, paid for by their respective universities.

But Wojcicki, who calls last Thursday "a very sad day for high energy physics and for our country", estimates that a useful US involvement in the LHC would mean spending at least \$500–\$700 million on a detector for the planned European accelerator.

He is deeply sceptical that such a sum would ever be approved by the US Congress "when we can't appropriate money for our own project". Wojcicki says he has considered resignation, but will stay to chair an HEPAP meeting in two weeks' time that will discuss how to respond to the termination of the discipline's dominant project.

Roy Schwitters, the director of the SSC, denies that recent allegations of mismanagement had been a decisive factor in the SSC's demise (see *Nature* 364, 92: 1993). "The essential line of attack was that curiosity-driven science is somehow frivolous, and a luxury we can no longer afford", ►

Around 900 of the 6,000 scientists who visit CERN each year, and one-third of its 1,000 staff scientists, are already deeply involved in this part of the project.

Individual member states are beginning to firm up their commitment. Earlier this month, for example, the Dutch government became the first to commit funds for the construction of one of the detectors, announcing that it had agreed to provide SFr12.5 million for this purpose.

A detailed costing of the collider will be presented to the CERN council when it meets in December. But there could be a problem if, as seems likely, this costing exceeds Rubbia's original estimates. If it does, member countries may be asked to increase their normal CERN contributions to cover it.

Government will then have a few months to consider their options before a decision is made at a special CERN council meeting planned for next April. But with so many CERN countries looking for ways to save money in their research budgets — and elsewhere — the US decision could become an excuse for some to revise their views about Europe's financial (and scientific) priorities.

Hermann Strub, for example, head of Germany's delegation to CERN, says that lobby groups within the scientific ►

SSC: End of an era

► he says. Schwitters says that his top priority for now is looking after the 2,000 staff at the SSC laboratory. "Everyone feels a sense of commitment to capturing what we have done", he says. Management at CERN and the SSC are already conferring on how the data can be captured and built upon.

The budget bill due to be passed by Congress this week includes \$640 million for termination costs at SSC in the coming year, the amount that was to have been spent continuing with the project. This money is supposed to cover restoration of the huge site to its previous state and 90 days' severance plus relocation expenses for the staff, as well as outstanding claims by hundreds of subcontractors.

But Brown says that, once legal settlements with the state of Texas and other likely litigants are included, two years of funding at that level will be needed to wind up the project. That will take the total cost of the debacle to more than \$3 billion. "It is very difficult for me to perceive why people in their right minds would pursue this course of action", says Brown.

Language in the budget bill requires O'Leary to prepare a report by next July on how to "maximize the value of the investment that has been made". Schwitters says his preliminary understanding is that land restoration will require only that tunnel entrances be closed and made safe, and not that the 15 miles of tunnel already dug be filled in again.

Scientific opponents of the collider were keeping a low profile after the decision. Philip Anderson, professor of particle physics at Princeton University, who has testified in Congress against the project, said he derived no satisfaction from its demise. He declined to comment further; "anything else I say now will be taken as gloating", he says.

But Jocelyn Hong, director of the Washington lobby group OOPS! (Organizations Opposed to the SSC), said the decision would benefit scientists on smaller projects in other disciplines, and noted the generosity of the SSC severance terms. "They get 90 days' money", she says. "No-one else gets that."

Colin Macilwain

SSC: European reaction

► community are now certain to campaign for the money to be diverted away from the LHC towards other scientific ends. Others argue that the existence of the SSC made it easier to argue the case for the LHC as a European rival.

On the other hand, the SSC cancellation could also provide an additional impetus for Europe to complete a scientific investigation that is already agreed to be the next logical step for particle physics.

Frank Close, head of theoretical physics at the United Kingdom's Rutherford Laboratory, describes the LHC as now "the only

Japan fears negative impact on domestic spending

Tokyo. The decision by the US Congress to terminate funding for the Superconducting Super Collider (SSC) could have a negative impact on funding for science in Japan, including Japanese participation in the European Laboratory for Particle Physics (CERN).

Since 1990, the United States has been putting pressure on Japan to contribute about \$1.5 billion to the construction costs of the SSC, a move resented by government officials and scientists in Tokyo.

But the pressure has also helped to make science a political issue in Japan, and has drawn attention to deficiencies in the research system, and in particular the poor state of its universities. Indeed, some Japanese academics fear that cancellation of the SSC could kill attempts to expand the government's budget for science.

Cancellation of the SSC could have a negative impact on negotiations with CERN, says a physicist at Japan's National High Energy Physics Laboratory (KEK), as US pressure for SSC support might have created a new funding mechanism that could

include a "small channel" to support Japanese participation in CERN. But not everyone shares this view; a university professor involved in CERN negotiations says it is not clear what effect the US decision will have.

But even if the SSC decision helps to clear the way for more decisive negotiations over CERN, there is no guarantee that the Japanese government will have any money to offer. "According to the newspapers many Japanese scientists are quite happy [about SSC decision]", says one science adviser to the Japanese government. "But I worry that it will have a bad effect on the expansion of budgets for basic science."

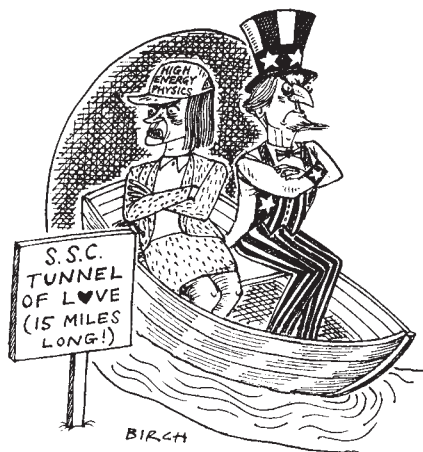
This adviser complains that Japanese politicians do not understand the difference between the SSC and other high-energy physics projects. As a result, he says, it may now become more difficult to win full funding for KEK's plans for a B-meson factory (see *Nature* 380, 330; 1993) let alone for CERN, because politicians and the Ministry of Finance are very reluctant to fund new projects in a time of recession.

Japanese supporters of the SSC are trying to make the best of last week's decision. "We are surprised and very disappointed", says Takahiko Kondo of KEK, leader of Japanese participation in the SSC, adding that Japan has "learned a lot" from participation in preparations for the SSC.

Funding from the US-Japan programme on high-energy physics has allowed Japanese scientists to conduct "risky research" that they were not able to do when building the TRISTAN electron-positron collider at KEK. As a result, they have developed new superconducting magnets and integrated circuits that will be patented, and will find uses elsewhere in high-energy physics.

Kondo predicts that international competition in B-meson physics will heat up as "US refugees" from the SSC turn to B-meson physics. But he worries that European governments may "cool" their support of CERN and that this could have an adverse world-wide effect on high-energy physics.

David Swinbanks



Mayor nominated for second UNESCO term

Paris. The executive board of the United Nations Educational, Scientific and Cultural Organisation (UNESCO) last week nominated Federico Mayor, the Spanish biochemist who is UNESCO's current director general, as its candidate for a second six-year term.

Mayor will be formally elected by the UNESCO general conference on 6 November. □

German opposition may delay Framework

Brussels. The European Communities (EC)'s next five-year Framework research programme, scheduled to start in 1994, may be delayed for more than a year because of Germany's wish to see a significant cut in the European Commission's proposed research budget of ECU13.1 billion (US\$15 million).

Elections to the European Parliament will be held next June. As a result, the EC has only a short window in which to complete the complex approval procedure for its fourth Framework programme. To meet the necessary deadlines, the council of research ministers must approve the budget unanimously at its next meeting in December.

If Germany, which is said to be seeking a figure as low as ECU8 billion, does not budge, the council will be forced to approve a large reduction in the budget. And if that happens, the European Parliament is likely to use the right of veto it was given by the Maastricht agreement, which comes into force at the beginning of November. The whole approval process would then have to be rerun in 1995.

Both France and the United Kingdom also challenge the size of the budget. But there is confidence in Brussels that both will be prepared to accept changes in the content of the Framework programme, and perhaps a slight reduction in the budget,

rather than risk a delay.

Belgium, in its role as current president of the Council of Ministers, is trying to find a compromise between the Commission, the Parliament and the Council. One possibility being considered would be to freeze about ECU1 billion of the budget.

Germany's science minister, Paul Kruger, is said to be keen to reach agreement with his colleagues. But the finance minister, Theo Waigel, may not let him. The cost of reunification has forced Germany to cut its domestic spending on science, making it politically difficult to agree to the proposed ECU6.2 billion increase in funding for Framework over the next five years.

Indeed, all three of the EC's three largest paymasters are reluctant to fund research controlled by Brussels when they are having to reduce spending on domestic research which they control directly. This reluctance is further fuelled by the continuing belief that the quality of many joint research projects is lower than that in their

own national programmes.

Claude Desama, the president of the Parliament's energy, research and technology committee, says he is "cautious" about the proposal to freeze part of the Framework budget. He argues that the increase in the budget being sought by the Parliament is not as large as it appears, as much of it covers the transfer of other EC research projects into Framework under the Maastricht Treaty.

Desama admits that a reduction in the budget request might allow Germany to save face at home. But he says that it would not save money, as any savings on research would be redeployed to other EC activities, such as regional development.

Fusion research could be one victim of any large reduction of the Framework budget. Funding for fusion has been almost doubled in the proposed programme, to ECU980 million. One parliament official says that for the first time many people are "thinking the unthinkable", namely abandoning the premium being paid to increase the chances of the International Thermonuclear Experimental Reactor being built in Europe.

The cost of the fusion programme was expected to go unquestioned until the end of the decade. But if the Framework budget is pruned, there will be pressure "to take it out of the shopping basket, and buy other goodies", he says.

Declan Butler



Desama: defends joint research spending.

'Precompetitive' research heads for the history books

Brussels. Pressure is mounting on the European Communities (EC) to abandon the rule specifying that research they support must be "precompetitive", that is, far enough from the market to let companies collaborate without giving away trade secrets.

The rule was introduced by Etienne Davignon, then research and industry commissioner, in 1984. His goal was to stimulate both industrial recovery and the single market by encouraging national champions to work together. Cooperation in research became the means of achieving this goal; and 'precompetitive' research was seen as a solution to the reluctance of industrial competitors to collaborate.

But Davignon, speaking at the European Science Summit organized by the European Parliament in Brussels earlier this month, suggested that the time has come to rethink the relevance of the rule to EC research. The policies he introduced may have encouraged European companies to merge into larger multinationals; but industrial and economic needs have changed.

The failure of EC research to close the technology gap with the United States and Japan has once again prompted a reflection on how, if at all, public money should be

used to promote innovation. Simultaneously, the Maastricht agreement has extended the goals of EC-funded research from merely "strengthening the scientific and technological basis of European industry" to promoting cooperation on virtually any research considered valuable to Europe.

One result is likely to be the disappearance of the term 'precompetitive'. "It's time to whip away the precompetitive figleaf", says a staff member of the Parliament's energy, research and technology committee.

If the research is first class, companies will want to do it on their own, he says. Most precompetitive research projects would never be done by companies if the EC was not paying half of the costs.

Many now feel that the notion of precompetitive research is based on an obsolete, linear model of innovation. In contrast, the European Parliament, among others, also wants EC funding to be reformulated on the basis of a 'systems' model of the innovation process.

For example, members of parliament say the EC would get more for its money in terms of boosting competitiveness by increasing support for both fundamental research and technology transfer. The European Com-

mission has already proposed increasing funding for technology transfer in the fourth Framework programme to ECU600 million — more than 5 per cent of the total. But many want it to spend much more.

Rolf Linkohr, rapporteur of the parliamentary committee, has also proposed that the EC puts more money into funding contract research organizations to disseminate and develop the results of the research programmes it finances.

A more controversial suggestion is that the EC should also fund applied research on the basis of its technological merit, irrespective of its distance from the market. But any such move would raise fears among free-marketeers that the research funds might become used to prop up flagging European companies.

In a report last month to François Fillon, the French minister of higher education and research, Jean-Pierre Chevillot, the vice-president of the French Conseil de la Recherche et de la Technologie, proposed that the EC should stop supporting precompetitive research and target its funding towards "partnerships" between public laboratories and private companies instead.

Declan Butler

Waldegrave defends 'foresight' exercise

London. William Waldegrave, Britain's science minister, has refuted claims that moves by the Office of Science and Technology (OST) to devise a strategy for choosing research priorities through a programme of "technological foresight" are running into difficulties.

But the OST seems to have implicitly accepted some of the criticisms of the programme that have been made in recent weeks. Speaking at a meeting at the Royal Society in London on Monday evening, for example, Waldegrave emphasized the need for research priorities to be determined by market needs, not merely by the 'technology push' that many feel characterized some early drafts of the programme produced by the OST.

He also acknowledged that pressures to achieve short-term results for the programme could undermine its long-term effectiveness. Waldegrave said that the truth was "somewhere in between" those who wanted quick action, and others who argued that the OST was already moving too fast.

The technology foresight programme is central to government's efforts to restructure its support for science and technology, as outlined in the recent white paper (policy document) *Realising our Potential*. In particular, it is being seen as the main device for imposing industrial priorities on the science base, for example by making research councils aware of those fields of science most likely to obtain government support.

A special committee, chaired by Bill Stewart, the government's chief scientific

adviser, has been set up to oversee the programme. And two outside consultants, the consulting firm Segal Quince Wicksteed of Cambridge and the Programme of Policy Research in Engineering Science and Technology (PREST) at the University of Manchester have been appointed to organize



Waldegrave: seeking a new partnership.

detailed surveys of expert opinion using the Delphi techniques pioneered in both the US and Japan.

In an attempt to draw attention to the exercise — and also to drum up support for it from a somewhat sceptical research and industrial commu-

nity — a series of seminars are being organized throughout the country.

Some of the initial scepticism has been dispelled by the seminars, particularly those which have included presentations from industrial research managers emphasizing the value of forecasting techniques in their own companies. A meeting at the Confederation of British Industry last week, for example, heard precisely such a pitch from John Taylor, research director for Hewlett Packard.

But the meetings have also provided a platform for expressions of concern over the way that the programme is being run so far. One is a perceived danger that it may be

taken over by the Whitehall civil servants. Evidence of this, it has been claimed, is the way that the OST produced a list of suggested technological priority areas, even before the consultation process had properly begun.

There are also widespread fears that the OST, under strong pressure from the Treasury to impose priorities on the science budget, may be trying to run before it can walk. "Why not concentrate on one or two areas first, and make sure you get those right", says one critic.

OST officials say they are well aware of each of these dangers, and are taking steps to minimize them. On the question of communication, for example, Waldegrave says a key function of the seminars — and indeed of the whole programme — is to help build a partnership between industry, academia and government.

OST's response has gone some way to placating its critics. "The whole project seemed to be in something of a mess a few weeks ago, but since then things have been picking up a bit", says Sir Mark Richmond, chairman of the Science and Engineering Research Council. But a high degree of scepticism remains.

Fiona Steele, of the Confederation of British Industry's (CBI) innovation unit, acknowledges that the OST is now listening to the CBI's concerns. But she says that there is still a long way to go before the confederation is satisfied that the programme is on track. "The proof of the pudding will be in the eating", says Steele. **David Dickson**

South African cyclotron faces uncertain future

Cape Town. South Africa last week became only the fifth country in the world — after Sweden, Japan, Russia and the United States — to offer high-energy proton therapy to cancer patients when the first such treatment was given to a patient at the country's National Accelerator Centre (NAC).

But controversy continues to dog the centre as questions are raised about the high proportion of the national research budget being absorbed by a facility now being used primarily for medical purposes. And the new government, which is due to be elected in April and which is likely to reassess research priorities, could well decide to withdraw its support.

Roger Jardine, who holds the science and technology portfolio in the African National Congress's Department of Economic Planning, says that the organization is not keen to support "big science" through projects intended to "glorify the state" as the National Party government has in the past.

The NAC is administered by the coun-

try's main science funding agency, the Foundation for Research Development (FRD). Its running costs are R34 million (US\$10 million) a year, a figure that represents just under 30 per cent of the foundation's operating budget.

In the past, this level of expenditure has been justified by foundation officials on the grounds that funding for the accelerator centre is earmarked by the cabinet, and could not be unilaterally reallocated to other scientific priorities.

The major facility at NAC is a variable energy cyclotron capable of producing 200-MeV protons (see *Nature* **327**, 271; 1987). Since the cyclotron became operational six years ago, the proportion of beam time used for nuclear physics has fallen to 35 per cent, partly because the energy range at which it operates is of relatively low scientific interest; the rest of the time is used for neutron therapy, radiobiology research and the production of short-lived diagnostic isotopes.

Treatment at NAC is provided free of

charge. But the number of patients who have received neutron therapy at the centre, for salivary gland and advanced breast tumours, is small, averaging about a hundred a year for the past five years.

Proton therapy is prescribed for tumours close to the spinal cord or optic nerve, and those in or near the brain, as it does not affect surrounding healthy tissue. Its advent should increase the number of patients treated by the centre. NAC receives only a small private income from the sale of isotopes: R670,000 last year.

If the facility fails to win the support of the new government, one option would be privatization. This could require a controversial shake-up among its personnel. The centre's director, Daan Reitman, revealed last week that well over half of its annual budget is needed for staffing costs. With a total of 200 employees, the average cost per employee is R100,000, an extremely high figure by South African standards.

Michael Cherry

Japan to study damage from Russian dumping

Tokyo. The uproar over Russia's dumping of low-level nuclear waste in the Japan Sea has drawn attention to a serious gap in attempts by Japan's marine scientists to monitor the possible contamination of a region in which thousands of tons of nuclear waste have been dumped by the former Soviet Union over the past 40 years.

It has also persuaded the two governments to bring forward the first meeting of a joint working group of specialists to discuss a survey of the dumping area which, ironically, had already been set up by the Russian president, Boris Yeltsin, during a visit to Tokyo earlier this month.

The disclosure by the environmental organization Greenpeace that Russia had resumed dumping in the Japan Sea occurred only days after Yeltsin signed a declaration in Tokyo expressing "serious concern" about such ocean dumping. Vehement protests from Japan and South Korea, backed by the United States, have already forced Russia to cancel plans to dump a further 800 tons of liquid waste.

But the 900 tons of low-level waste dumped by the Russian Navy under the gaze of Greenpeace cameras is relatively small



Japanese agencies visited numerous offshore survey sites earlier this year.

compared to the activities of the former Soviet Union. In March, Russia revealed that, since 1959, more than 80,000 tons of low-level liquid waste, more than 5,000 containers of solid low-level waste, 34 ships packed full of waste, and 2 nuclear reactors (with fuel removed) had all been dumped in the Japan Sea in an area a few hundred kilometres south of the Russian port of Vladivostok.

Between April and June, the Japanese Meteorological Agency, the Fisheries Agency, the Maritime Safety Agency, and the Japan Marine Science and Technology Centre (JAMSTEC) all sent ships to analyse sea water, deep-sea sediments and fish for various radionuclides in a number of

locations (see map).

No unusual levels of radioactivity were found, and at the end of August, Japan's Science and Technology Agency (STA) announced that "the results of the survey so far show that this ocean dumping is not something that has an effect on the health of our people".

But the conclusions are of limited value, as none of the surveys was able to collect samples in the actual dumping area, which lies within Russia's 200-mile exclusive economic zone. Investigations have been confined to areas off Japan's coast, tens to hundreds of kilometres distant from the dump site.

Japanese marine scientists need samples from the dumping region to reach definitive conclusions, and Japan has been pressing Russia to allow this. The joint meeting agreed to by Yeltsin, which will be held this week, is seen as a step in this direction.

An STA official says Japan hopes to conduct a full survey of the dumping region "this year or next". And, in addition to budgets already allocated this fiscal year, the agency has requested ¥306 million (about \$3 million) for further surveys next fiscal year, which begins on 1 April 1994.

Russia itself is hoping to persuade Japan to provide financial support for the construction of a land-based facility for waste disposal. Japan is considering allocating part of an agreed \$100 million in grants for dismantling nuclear weapons for this purpose. But, even if the land-based facility is built and sea dumping stops, surveys will have to be continued for many years to monitor the effects of waste disposal over the past four decades. **David Swinbanks**

Childhood leukaemias remain unexplained

London. After a three-year investigation by the Health and Safety Executive (HSE), British health officials say they remain baffled by an outbreak of leukaemia among children of men living in the village of Seascale in Cumbria who worked at the nearby Sellafield nuclear plant.

The investigation confirmed findings first reported three years ago by the late Martin Gardner, then head of the Medical Research Council's Environmental Epidemiology Unit at Southampton General Hospital, that an observed excess of leukaemias appeared to be correlated with the radiation dosages received by fathers from Seascale working at Sellafield.

But the report also confirms that the excess of cancers could have been caused by a process of "population mixing", proposed by Leo Kinlen of the Radcliffe Infirmary in Oxford, in which the risk of childhood cancers is claimed to increase when people come to live in a geographically isolated area from other parts of the country.

Both the nuclear industry and its critics have drawn support for their arguments from the HSE report. The first claims that it confirms the lack of proof that the nuclear plant was to blame for the cancers; the second that it does nothing to undermine this hypothesis and strengthens the case that the plant could be to blame.

Dumping boosts case for total ban

Tokyo. Revelations that Russia has been continuing to dump nuclear waste in the Japan Sea (see above) may force Japan to support a total ban on the ocean dumping of nuclear waste at a meeting in London early next month.

The London Convention on ocean dumping was established in 1975 to prevent contamination of the sea by nuclear waste. Japan became a party to the convention in 1980. The convention bans all dumping of high-level waste. And in 1983 and 1985, many of the countries belonging to the convention also agreed to a voluntary ban on dumping low-level waste.

Both Japan and Russia abstained from voting in 1985, and have therefore not committed themselves to the voluntary

ban. But René Coenen, senior technical officer for the London Convention at the International Maritime Organization in London, says that, whatever the legal position, the recent dumping by Russia goes "against the spirit of the convention".

At a meeting in London on 8 November, seven proposals, ranging from a complete ban on dumping of radioactive waste to resumption of some dumping, will be discussed.

Until now, Japan has been undecided on how to vote because its nuclear power industry has been urging the government to retain an option on dumping. But following Japan's protests about Russia's activities in the Japan Sea, it will be hard for Japan not to vote for a permanent ban. **D.S.**

German university faces ban on animal tests in class

Munich. The department of zoology at the University of Marburg in Germany has been told by the local Hessen government to stop using experiments on anaesthetized rats as part of an undergraduate teaching course in animal physiology.

The Hessen government, which is led by the Green Party, says that it eventually wants to ban the use of animals from all undergraduate courses. But the university is appealing against the injunction, claiming that it is neither in the interests of students nor in the long-term interests of society.

The row began five years ago, and has centred on a single experiment to study the intestinal absorption of glucose in an anaesthetized rat, the only animal experiment now left during the whole of the course.

The main aim of the experiment, according to Gerald Heldmaier, the head of the zoology department, is to teach students about the control of anaesthesia, and introduce them to simple surgical techniques. In each case, students must pass a test on the relevant theory in advance.

Germany's animal protection law states that animal experiments may be carried out for teaching purposes only where no viable alternative exists. Students at the university argue that the experiment could be replaced by a film which they themselves made four years ago. But Heldmaier, who has wide support among his colleagues, maintains that practical experience is an essential component of his course.

Last February, the research minister of Hessen sent a letter to the university stating that its long-term aim is to ban the use of both whole animals and parts of animals in

teaching. In reply, 63 professors at the university signed an appeal saying that such a move would lower the standards of education and endanger the quality of future researchers.

According to Heldmaier, discussions that had been promised between the university and the *Land* government failed to materialize. And this summer, the students stepped up their protests by taking out a private legal action accusing Heldmaier of coercion in making the experiment a course requirement, and of violating the country's animal protection act.

Over the past few years, about 20 students out of an average of 100 taking the animal physiology course have refused to carry out the experiment. On graduation, they have been awarded a 'qualified' certificate, indicating that they did not complete this part of the course.

Irvana Weber, a spokesperson for the students, says that their main criticism is that the experiment is mandatory for all students. Heldmaier replies that only those majoring in zoology are required to conduct it; others majoring in disciplines that would not lead to careers working with animals can choose whether or not to do so.

Support for Heldmaier has come from the German Physiological Society. Earlier this year it adopted a resolution, signed by nearly all of Germany's physiology professors, stating that the nature of their subject requires students to complete practical experiments on animals or parts of animals. Last week the society wrote to Heldmaier saying they believed this also to be true of zoology.

Alison Abbott

Cloning of human embryos draws fire from critics

Washington. Critics of research announced earlier this month which has succeeded in cloning human embryo cells for the first time have pledged to block any repetition of it in court and in Congress.

The successful cloning of the embryo cells — which surfaced discreetly at a low-key meeting of fertility researchers on 13 October — is certain to reopen a fierce debate about the ethical boundaries of human embryo research.

Science activist Jeremy Rifkin described the research, which applied techniques commonly used to clone animal embryos, as "very troubling". He accused the researchers involved — Robert Stillman and Jerry Hall of the *in vitro* fertilization department at George Washington University Medical Center in Washington — of doing work "of no redeeming value" just because "they wanted to see if it would work".

Stillman says that research, which was approved by the Medical Center's ethics committee, involved only genetically abnormal embryos which had been penetrated by more than one sperm, and could not have become viable.

He says that the results "indicate that the process has the potential for reducing risks and costs as well as enhancing success rates for infertile couples who desire children." Stillman adds that the experiment has no clinical value for now, but should be regarded as a "small step in an on-going scientific process".

In the experiment, 17 embryos consisting of a few cells each were divided into 48 single cells, covered in a coating developed by Hall and placed in a culture. The single cells divided an average of three times before dying; some of them survived to divide five times, into 32-cell embryos.

The research was first reported at a joint meeting of the American Fertility Society and the Canadian Fertility and Andrology Society. It was part of a series of studies aimed at helping infertile couples. But Hall and Stillman say that the research which produced the cloned embryos has now ceased.

Although the work at GWU is not funded by US federal agencies, Rifkin's lobby group, the Foundation on Economic Trends, has filed a petition asking the National Institutes of Health (NIH) to halt all funding of "institutions working on human cloning."

Such a sanction would effectively bar even privately-funded work from taking place in any US university. If NIH refuses to act, Rifkin has promised to follow through with legal action. He has also called on members of Congress to make all human cloning illegal in the United States.

Colin Macilwain

UK scientists protest lab privatizations

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A large brain carrying a "for sale" notice was paraded outside the House of Commons last week by members of the Institute of Professional Managers and Specialists in protest at the government's policy of selling off many of its research laboratories to the private sector. The union, which represents scientists working in government and research council laboratories, says that over 34,000 jobs are likely to be affected.

Stefano Cagnoni

Indian tree offers nuclear waste treatment

New Delhi. A herbal product claimed to be able to mop up uranium and other long-lived isotopes from nuclear waste has been developed by Indian scientists applying modern biological techniques to a traditional method for cleaning dirty water.

The product has been isolated from the seeds of *Strychnos potatorum*, a tree found growing in forests in the state of Andhra Pradesh. It is also said to be capable of removing cadmium, mercury and other toxic heavy metals from factory effluent, and to have potential uses in the mining industry, as it can bind to metals such as gold, silver, cobalt, copper and nickel.

For centuries, indigenous people living in the state have been aware of the seeds' ability to precipitate fine particles into larger masses. Even today the powdered seed is used to transform muddy water in streams and wells into clean drinking water.

Realizing that the seeds of the little-known tree, which grows only in India, Burma and Sri Lanka, might have many industrial applications, the state's Girijan

Cooperative Corporation (GCC) hired a team of scientists two years ago to conduct research on their remarkable properties.

"All that we did was to isolate the substance responsible for the property which the tribals already knew", says Durga Prasad, a biochemist and leader of the research team. The product does not yet have a name; the scientists refer to it simply as a bioflocculant.

According to Prasad, the ease with which the substance binds with a variety of metals is dependent on the pH of the reaction medium. The binding efficiency has been found to be high; one mole of bioflocculant can apparently absorb between 20 and 25 moles of metal. Prasad says that the ability of the seeds to bind metals appears to result from the presence of certain proteins.

This has been confirmed by Clement Furlong, professor of biochemistry at the University of Washington in Seattle, who is carrying out a joint research project with GCC to study the substance's properties.

"These proteins are apparently unique", says Furlong. He said his collaboration has

two aims: to purify the proteins still further, in order to improve their specificity and metal binding ability; and to clone the genes coding for the proteins, so that they can be manufactured using the techniques of biotechnology.

The herbal product has a very high affinity for uranium, and might therefore be used for removing traces of this metal from reactor effluent, a characteristic that has already attracted the attention of the International Atomic Energy Agency (IAEA).

K. Mahadev Rao, a senior official of the IAEA, says that the agency is exploring its possible use in the manufacture of cobalt-free steel for the nuclear industry. Cobalt becomes radioactive after it has been irradiated by neutrons, but there is no way of making steel free of contamination by cobalt.

According to Rao, the ability of the bioflocculant to bind with actinides means that another possible use is in nuclear waste disposal, particularly because the selective removal of actinides makes the remaining nuclear waste easier to handle.

Metal binding biochemicals are not new. But Prasad claims this is the first time that a biological substance has been isolated from a plant capable of binding with a wide range of metals.

K. S. Jayaraman



***Strychnos potatorum*: its seeds can bind to toxic metals.**

S. Korea pauses on science reforms

Seoul. President Young-Sam Kim of South Korea has called a temporary halt to negotiations over a planned reorganization of the countries' science-related ministries because of disagreement between them about the form that the reorganization should take.

South Korea has changed dramatically over the past three decades from a poor agricultural society to an industrialized nation with a developed research infrastructure. But the growing involvement of different ministries and agencies in research has put a strain on the distribution of responsibilities, and made reorganization necessary.

Shortly after coming to power earlier this year, Kim promised such a reorganization, and it had been expected that his plans for achieving it would be presented to parliament this month (see *Nature* 364, 384; 1993). But this move has now been postponed to allow further time for discussion.

Some of the more powerful ministries in Seoul, such as those of trade and industry and communications, want to reduce the powers of the relatively young Ministry of Science and Technology (MOST). They would like to see MOST take on more of a purely coordinating role as they themselves become increasingly involved in supporting research and development.

These ministries would, for example, like greater responsibility for the budgets of individual research institutes. A step in this direction was the recent transfer to the Electronics and Telecommunications Research Institute, previously funded by MOST, to the ministry of communications.

Others, however, would prefer to see a strengthening of the Presidential Council on Science and Technology, whose members include the Minister of Science and Technology. This could also lead to a stronger role for MOST, and in particular its Science and Technology Policy Institute.

And there are some who argue that the main need is to boost the priority given to science and technology by the Economic Planning Board (EPB), Korea's equivalent to Japan's powerful Ministry of Finance. This could, for example, be done through the creation of a new section in the EPB.

Ever since the proposed reorganization was announced, different ministries and agencies have been giving their own, often opposing, views to the president. This process has now been halted to give the president and his advisers more time to consider the implications of the different options.

As part of this process, the presidential council last week held the first of a series of meetings with foreign science policy specialists to see what can be learnt from the experiences of other advanced industrialized nations, in particular Japan.

Other advisory bodies, such as a restructuring committee and economic advisers to the president, are also expected to present their views on what changes should be introduced. And the final plan could contain elements of all the various alternatives now being put forward. But there is still uncertainty over how much time should be spent on building consensus before changes are put into effect.

David Swinbanks

Paradigm lost

SIR — Christopher Exley's wistful reflections on the redefining of basic science (*Nature* 364, 276; 1993) deserve serious consideration. His letter reminds us clearly of the chronic frustration and demoralization experienced by academics engaged in the pursuit of knowledge for its own sake. What seems to me to strike at the heart of this despondency is not so much the lack of funds and poor professional standing — although these clearly are important factors — but the emptiness that accompanies the realization that the incessant privatization of basic science is often at the expense of individual creativity. It is as if the very essence of academic research has become a stigma.

I can only assume that the advocates of current policies would like to disarm fundamental research, and its inherent unpredictability, in favour of a virtual reality based on scientific clairvoyance. The fundamental flaw in this approach is that the personal needs and aspirations of talented people to express their scientific creativity are incommensurate in the long term with preordained visions of technological exploration. In a similar vein, the current move towards the straitjacketing of postgraduate training in the United Kingdom will deny young people the opportunity to experience the underlying aesthetic quality of scientific discovery. It is my contention that, without this experience, new advances in science will become curtailed because such a policy offers little that is innovative within the sphere of the individual.

The freedom to follow up one's curiosity, to think laterally, to (dare I say it) "play", are important prerequisites for basic research that cannot be discounted if new knowledge is to be realized. The fact that this is a paradigm lost will not only serve to undermine the future exploitation of science for economic gain but is a serious unjustified impoverishment of the cultural vitality of industrialized nations.

Stephen Mann

*University of Bath,
School of Chemistry,
Claverton Down,
Bath BA2 7AY, UK*

SIR — Philip Siekevitz, in his letter headed "Love or money?" (*Nature* 364, 477; 1993), notes the growing habit of pharmaceutical companies who buy the rights to exploit the scientific findings of independent research institutes, many of which have long received substantial support from public monies, rather than fund basic research in their own organizations. He also eloquently bemoans the continued privatization of science by researchers who see in their own work the glimmer of personal profit and then goes

on to ask "What has become of the love of science, of research, of simple satisfaction with a job well done?"

What has become of those with these attributes? The sad fact is that many are no longer career scientists because of the lack of continuing public support for their science; some of us spend increasing amounts of potential research time competing for the scarce public dollars still available; and the remainder, as Siekevitz notes, sell themselves to the highest bidder. In my experience, most of these latter individuals are not driven by greed, but rather, as we all are, by the desire to continue the personal quest for basic answers to some of Nature's most important questions. However, as public support of basic science continues to shrink in real dollars, more and more of the existing science enterprise is, in effect, put on the block. After all, pharmaceutical companies cannot buy what has not already been devalued, abandoned and, in effect, put up for sale.

I recently returned from a meeting sponsored by the US National Science Foundation (NSF) for directors of its

Research Experiences for Undergraduates Program. This programme, designed and funded by NSF in the hope that it will have a positive impact on career choices in science, annually provides hundreds of college students with the opportunity to participate directly in the research enterprise by placing them in laboratories in universities and research institutes across the nation.

Siekevitz would be pleased to know that love of science continues to be a strong motivator for most of these students. However, many of them say that they will not pursue this path because of the growing perception that a career in basic science is not an appropriate choice for those who also need the money.

Tragically, for most of us, it has never been a question of "love or money?", but like many other areas of human enterprise it remains, love and money.

Robert J. O'Connell

*Worcester Foundation
for Experimental Biology,
222 Maple Avenue,
Shrewsbury,
Massachusetts 01545, USA*

Non Angli . . .

SIR — Your leading article (*Nature* 364, 745; 1993) makes a number of important and valid points about the decline in the numbers leaving school with the qualifications expected for the study of science and engineering in British universities. Inconsistency and the promulgation of incompatible policies are, however, characteristics that we have grown accustomed to expect from this administration.

One of the issues that you discuss at length is the inadequacy of the A-level system. Such problems are not, of course, "British" as you state, but particular to England and Wales. How many times do London-based journalists have to be reminded that here in Scotland — a part of Britain — there is a different education system, with five 'highers' being the normal programme of study for entrance to universities?

The system of 'highers', with clear weaknesses of its own, attempts to address the central concern with A-levels expressed by those in England and Wales. Students with 'highers' have been known to gain entrance to universities south of the border, as well as in Scotland, although in general the Scottish university institutions adjust to the different knowledge base of such students through a longer course of study.

Paul Trayhurn

*Division of Biochemical Sciences,
Rowett Research Institute,
Bucksburn,
Aberdeen AB2 9SB,
Scotland, UK*

Fair shares

SIR — In recent issues of *Nature* there has been increasing comment about ways in which the current falling levels of postgraduate students could be halted. From my time as a student I feel that one of the most important parts of being a research student is to feel that one is an 'accepted' scientist. One of the best ways to achieve this is by presentation of the student's work at scientific congresses. I feel, however, that the old idea of a research group attending as a unit is rapidly being replaced by either the principal researcher or a single student attending because of the exorbitantly high costs now being charged by professional congress agencies. For example, at a forthcoming congress in my research field, it will cost DM1,600 per person (£640) to attend, at the cheapest rates and not using air travel. That comes to more than £2,500 for our small group. This is a sizeable amount and restricts attendance to one meeting a year or one delegate per congress, often the principal investigator. This restriction equally applies to researchers from Eastern Europe, or when air travel is involved.

Perhaps it is time that congresses were once again held at universities rather than city centre hotels in popular cities or resorts with the subsequent higher costs entailed?

William Brooks

*Institute for Surgical Research,
Ludwig-Maximilians-University,
Klinikum Grosshadern,
81366 Munich,
Germany*

The pathway to signal achievement

Sean E. Egan and Robert A. Weinberg

There has been a deluge of reports, in this journal and elsewhere, dealing with the subject of signal transduction and the Ras protein. Why all the fuss? The following is a News and Views article with a difference — an explanation for those not close to events, and a celebration of an achievement in biomedical research.

DURING this summer, an entire field of research passed a milestone. The field is cellular signal transduction featuring the apparently ubiquitous Ras protein. What has happened is that the main pieces have now been placed in a puzzle that has challenged researchers for more than a decade. The picture revealed is of an apparently universal signalling pathway through which many of the signals that direct nucleated cells to divide and differentiate are channelled. The pathway begins with cell-surface receptors and ends in the cell nucleus with proteins that regulate gene expression.

Signalling pathways within cells are formed by chains of intercommunicating proteins. Each protein component of a pathway integrates signals from upstream activators and passes them on to various downstream targets or effector proteins. The acquisition and subsequent release of signalling information is the process of signal transduction.

Individual transducing proteins work in a variety of ways. They may respond to acquired signals by amplifying them, damping them down and processing them in complex ways before they transmit them to downstream targets. This has the ring of electronic circuitry, and one day we will surely be able to depict these pathways in a form recognizable to the electrical engineer.

Regulation

Such pathways enable a cell to respond to external factors regulating cell division and differentiation. These pathways do so by integrating a variety of signals received at the cell surface. The signals enable a cell to judge the status of its cellular neighbours, its contact with substratum, and the presence of growth factors and hormones, before it decides to divide or differentiate. The interest in signalling pathways stems, in part, from the fact that when they go awry a cell may begin to multiply uncontrollably in a fashion that leads ultimately to cancer.

The dissection of the intracellular signalling pathways in control of cell growth and multiplication was stimulated by the discovery of genes that are responsible for forcing a cell to become cancerous. These oncogenes were initially discovered as the source of the cancer-causing powers of many RNA tumour viruses (retroviruses); later, they were found in mutated form in

many non-virally induced tumours including those of humans.

From the beginning, oncogenes were suspected to encode aberrantly functioning members of the cellular signalling cascades dedicated to the transmission of growth-regulating signals. By releasing signals in a deregulated fashion, oncogene-encoded proteins were presumed to activate portions of signalling cascades that lay downstream from them, resulting in a flood of growth-stimulatory signals inside the cell.

The proteins encoded by the cell's *ras* genes were presumed to be central to one of these cascades, and by the mid-1980s mutant forms of *ras* had been found in a wide variety of human tumours¹. These oncogenic versions differ from the normal version by the change of a single base of their nucleotide sequence. This change affects, in turn, the structure of the encoded Ras proteins, which carry an incorrect residue at one point in their amino-acid chains. But such discoveries, although provocative, revealed nothing about the position of the Ras proteins in cell signalling pathways. The two important issues — what activates the Ras proteins and, once activated, how they unload their downstream signals — remained a mystery.

The normal products of the *ras* gene are members of a large family of proteins^{2,3} known to bind guanine nucleotides. Like other members of this class, a Ras protein functions like a binary switch. In its OFF position, Ras binds the nucleoside diphosphate GDP and sits patiently at its usual location, the inner surface of the cell membrane, awaiting an incoming, activating signal. After receiving such a signal, it sheds its GDP, binds GTP (the triphosphate) and, for a brief moment of glory, exists in the ON state, releasing a pulse of growth-stimulating signals into the cell. Its lifetime in the ON position is guaranteed to be brief, as the bound GTP is rapidly converted to GDP by activation of the Ras protein's intrinsic GTPase activity. This conversion returns Ras back to its OFF state.

The oncogenic forms of Ras proteins found in many cancer cells work differently. They have greatly reduced GTPase activity inside the cell, and therefore do not get turned off; instead, they remain on for extended periods and flood the cell with growth stimulatory signals. The roots

of cancer could be traced with great precision to subtle changes in a particular molecular switch.

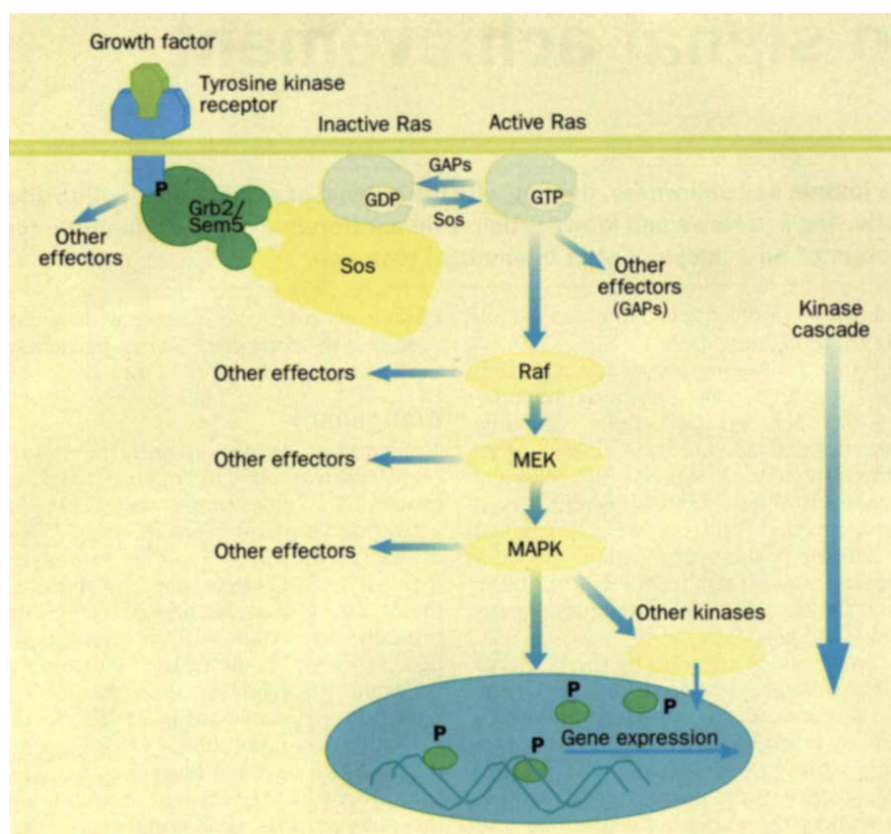
Malignancy

Yet another disruption perturbing Ras regulation was found to be connected with cancer. It is now known that GTPase-activating proteins (GAPs, see box) associate with Ras proteins and stimulate their intrinsic GTPase, thereby shutting them off⁴. Curiously, one of the GAP proteins was found to be present in defective form in the cells of individuals suffering from inborn susceptibility to type-1 neurofibromatosis (NF)⁵⁻⁹, a pre-malignant condition. Once again, the inability to shut down Ras activity led directly to malignancy. But for all the progress that was being made, the place of Ras in a larger signalling cascade was unknown.

During the 1980s, other oncogene products were characterized that, like Ras, could transform normal cells into tumour cells. This pot-pourri of oncoproteins displayed a wide variety of enzymatic activities¹⁰. Some encoded a protein tyrosine kinase, an enzyme that phosphorylates specific tyrosine side-chains of other protein molecules. Certain cell-surface receptor oncogenes were also found to carry such tyrosine kinase activity in their protein chains: a well-studied example was the receptor for epidermal growth factor. Other oncogenes encoded polypeptide growth factors. Still others encoded kinases able to phosphorylate specific serine and threonine amino-acid residues of target proteins. Each new report added to the feeling that this field of research was becoming a morass of complexity with no sign of an underlying order.

Developments over the past few months, however, have revealed how many of these oncoproteins participate in an evolutionarily conserved signalling pathway with Ras at its centre¹¹ (see figure). As more insight is gained into the workings of individual oncoproteins, the complexity is being replaced by an elegant picture of the Ras signalling pathway, which seems to be present (with slight variations) in eukaryotic organisms from yeast to man.

The universality of Ras signalling has meant that studies in many different organisms, some of which are susceptible



The Ras pathway (here stripped of adornment) which is central to growth control in organisms as diverse as humans, fruitflies, nematode worms and yeast. It runs from the cell surface to the nucleus, where genes are turned on or off in response to the incoming signal. The process starts with the binding of a growth factor, such as epidermal growth factor, to its tyrosine kinase receptor, resulting in autophosphorylation of the receptor's tyrosine residues. Controllers of Ras exchange factors, such as Grb2/Sem5, then come into operation. They recruit exchange factors (or activators, Sos being an example) to interact with Ras; the activators function as guanine-nucleotide-releasing agents which convert Ras-GDP, the inactive form, to active Ras-GTP. Downstream of Ras, a principal target is the Raf protein, also the product of an oncogene. Once activated, Raf phosphorylates a second kinase (MAP kinase kinase, or MEK). After mediation by other kinases, signals pass into the cell nucleus through phosphorylation of transcription factors that regulate gene expression. Not shown are other signalling routes that interact with this pathway, or details of branch or feedback points (see box).

to elegant genetic manipulation, have contributed greatly to its elucidation. Ras is used for a bewildering variety of purposes. In baker's yeast, *Saccharomyces cerevisiae*, the Ras pathway is involved in sensing nutrient conditions and controlling mitotic growth cycles^{12,13}. In these yeast cells, Ras receives signals from a guanine-nucleotide-exchange protein that encourages release of GDP and subse-

quent GTP binding¹²⁻¹⁴. There is a similar protein upstream of Ras in the fission yeast *Schizosaccharomyces pombe*¹³; here Ras controls the response to mating pheromone as well as meiosis, sporulation and cell shape^{13,15}. In *S. pombe*, the Ras protein, in turn, is upstream of a series of protein kinases^{13,15,16}.

In the fruitfly *Drosophila melanogaster*, genes controlling the development of the

complex eye and those specifying terminal structures of the early embryo specify a series of proteins that can be arrayed in a pathway that begins with a cell-surface tyrosine kinase receptor, passes through a Ras protein, and ends with a serine-threonine kinase^{17,18}. In the nematode worm *Caenorhabditis elegans*, genes effecting proper development of the vulva were similarly found to encode proteins that could be arrayed in a pathway¹⁹. Strikingly, the arrangement of the pathway in worms was very similar to that constructed for the fruitfly. A Ras protein sits squarely in the middle of each of these signalling chains.

Most recently, the protein made by the *raf* oncogene, long known for its ability to mimic some of the effects of Ras, has been found to embrace Ras²⁰⁻²⁴, forming a bimolecular complex. This level of physical intimacy usually implies that the two partners are talking to one another. In this case, Ras appears to be passing signals down to the Raf protein (but see box). On the upstream side, a number of distinct exchange factors (Sos in *Drosophila* is one example), also called guanine-nucleotide-releasing factors, are now known to activate Ras, doing so by forcing it to release its old GDP and take on a fresh GTP molecule²⁵⁻²⁸.

Connections

The story gets even better, for connections are being forged further upstream and downstream in the signalling cascade. Upstream, controllers of Ras exchange factors have now been found²⁹. Termed Grb2 by the protein biochemists^{30,31} and Sem5 by those interested in vulval development of the worm³², they embrace exchange factors controlling Ras. Grb2/Sem5 also binds to phosphate groups that tyrosine kinase receptors attach to their own tails when they are active — these decorations are only formed when the cell-surface receptors at the head of the stream receive signals from the extracellular space^{33,34}.

Other striking connections have been made on the downstream side. It turns out that the Raf protein, once activated, phosphorylates a second kinase, known variously as MAP kinase³⁵⁻³⁷ or

1. Bos, J. L. *Cancer Res.* **49**, 4682-4689 (1989).
2. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **348**, 125-132 (1990).
3. Lowy, D. R. & Willumsen, B. M. *A. Rev. Biochem.* **62**, 851-891 (1993).
4. Trahey, M. & McCormick, F. *Science* **238**, 542-545 (1987).
5. Viskochil, D. *et al.* *Cell* **62**, 187-192 (1990).
6. Cawthon, R. M. *et al.* *Cell* **62**, 193-201 (1990).
7. Wallace, M. R. *et al.* *Science* **249**, 181-186 (1990).
8. DeClue, J. E. *et al.* *Cell* **69**, 265-273 (1992).
9. Basu, T. *et al.* *Nature* **356**, 713-715 (1992).
10. Bishop, J. M. *Cell* **64**, 235-248 (1991).
11. Smith, M. R., DeGudicibus, S. J. & Stacey, D. W. *Nature* **320**, 540-543 (1986).
12. Broach, J. R. *Trends Genet.* **7**, 28-33 (1991).
13. Powers, S. *Sem. Cancer Biol.* **3**, 209-218 (1992).
14. Jones, S., Vignais, M. L. & Broach, J. R. *Molec. cell. Biol.* **11**, 2641-2646 (1991).
15. Egel, R., Nielsen, O. & Weiglun, D. *Trends Genet.* **6**, 369-373 (1990).

16. Errede, B. & Levin, D. E. *Curr. Opin. Cell Biol.* **5**, 254-260 (1993).
17. Greenwald, I. & Rubin, G. M. *Cell* **68**, 271-281 (1992).
18. Perimon, N. *Cell* **74**, 219-222 (1993).
19. Horvitz, H. R. & Sternberg, P. W. *Nature* **351**, 535-541 (1991).
20. Moodie, S. A., Willumsen, B. M., Weber, M. J. & Wolfman, A. *Science* **260**, 1658-1661 (1993).
21. Van Aelst, L., Barr, M., Marcus, S., Polverino, A. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6213-6217 (1993).
22. Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. *Cell* **74**, 205-214 (1993).
23. Zhang, X.-F. *et al.* *Nature* **364**, 308-313 (1993).
24. Warne, P. H., Vicinia, P. R. & Downward, J. *Nature* **364**, 352-355 (1993).
25. Feig, L. A. *Science* **260**, 767-768 (1993).
26. Simon, M. A., Bowtell, D. D. L., Dodson, G. S., Lavery, T. R. & Rubin, G. M. *Cell* **67**, 701-716 (1991).
27. Bonfini, L., Karlovich, C. A., Dasgupta, C. & Banerjee, U. *Science* **255**, 603-606 (1992).
28. Bowtell, D., Fu, P., Simon, M. & Senior, P. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6511-6515 (1992).

29. McCormick, F. *Nature* **363**, 15-16 (1993).
30. Lowenstein, E. J. *et al.* *Cell* **70**, 431-442 (1992).
31. Matuoka, K., Shibata, M., Yamakawa, A. & Takenawa, T. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9015-9019 (1992).
32. Clark, S. G., Stern, M. J. & Horvitz, H. R. *Nature* **356**, 340-344 (1992).
33. Schlessinger, J. & Ullrich, A. *Neuron* **9**, 383-391 (1992).
34. Pawson, T. & Gish, G. D. *Cell* **71**, 359-362 (1992).
35. Kyriakis, J. M. *et al.* *Nature* **358**, 417-421 (1992).
36. Howe, L. R. *et al.* *Cell* **71**, 335-342 (1992).
37. Hughes, D. A., Ashworth, A. & Marshall, C. J. *Nature* **364**, 349-352 (1993).
38. Crews, C. M., Alessandrini, A. & Erikson, R. L. *Science* **258**, 478-480 (1992).
39. Pelech, S. L. & Sanghera, J. S. *Trends biochem. Sci.* **17**, 233-238 (1992).
40. Gutman, A. & Waslylyk, B. *Trends Genet.* **7**, 49-54 (1991).
41. Hunter, T. & Karin, M. *Cell* **70**, 375-387 (1992).
42. Montminy, M. *Science* **261**, 1694-1695 (1993).

Cross-talk and versatility — the bigger picture

SIMPLICITY has emerged from complexity in unveiling the Ras pathway. But that pathway must also be considered in the context of other, equally fundamental cellular signalling cascades. The cross-talk between them yields subtle, finely tuned control.

Some of these other pathways employ low-molecular-weight, soluble compounds to transmit signals throughout the cell. These so-called second messengers include cyclic adenosine monophosphate (cAMP), calcium ions, diacylglycerol and phosphorylated forms of the sugar inositol. When produced at the plasma membrane, for example, they can activate specific protein kinases at distant sites in the cell. The protein kinase C (PKC) family contains examples of such target kinases and it has been shown that PKC- α efficiently phosphorylates and activates the Raf protein; so Raf integrates upstream signals from both Ras and PKC. From studies in yeast, it is also clear that Ras donates signals to many effectors or targets. In mammalian cells, the Raf protein may well be the most important growth-regulating effector, in that like Ras it is capable of transforming cells on its own. But the GAP proteins p120^{GAP} and p250^{NF} are also suspected to be Ras effector proteins regulating several different downstream functions.

There is versatility at the head of the pathway as well. When a polypeptide growth factor (such as epidermal growth factor) binds to its tyrosine kinase receptor, many different effector proteins are activated

through interaction with the autophosphorylated tail of the receptor; only one of these proteins is Grb2. Grb2 regulates the location and/or activity of several proteins; only one of these is Sos. Sos may itself have several functions, one of which is to activate Ras, and GTP-bound Ras in turn activates several effector systems including Raf. The Raf, MEK and MAPK proteins probably also have many targets.

In this way, genes are regulated in response to multiple signals at each point in the pathway and not just those emitted from the previous link in a linear chain. Indeed, the very fact that proteins such as the tyrosine kinase receptors, Ras and possibly Raf, have many targets may make them uniquely susceptible to oncogenic activation. Mutant alleles of these genes would then perturb entire signalling networks in ways that cannot be compensated for by other, less-powerful signal transducing components which lie on only one pathway.

As well as the Ras pathway described here, there are other independent oncogene-mediated pathways responsible for transmission of signals from the cell surface to the nucleus. They include the *TANI/int-3* oncogene pathway(s) which have been studied extensively in flies (*Notch*), worms (*lin-12*, *glp-1*) and frogs (*Xotch*); and the *Wnt* pathway in flies (*wingless*) and frogs (*Xwnt*). In addition, the *rel*, *bcl-3*, *lyt-10* oncogene signal transduction pathway has been well characterized in

flies (*toll*, *tube*, *pelle*, *cactus*, *dorsal*) and in the context of transcriptional regulation by the NF- κ B proteins.

Another potent class of growth-regulating signalling proteins is the steroid receptors, such as the *erbA* oncoprotein, which receive steroid-hormone signals and in turn alter gene expression. Moreover there are products of oncogenes (*db1* and *ect2*) and anti-oncogenes (*NF2*) which may be involved in cytoskeleton-mediated signal transduction — the *db1* and *ect2* proteins interact with Ras-related small GTPases, which are in turn important regulators of cell shape and motility, and *NF2* is similar to proteins also involved in control of cell structure.

The upshot of seeing the Ras pathway as part of this bigger picture is that although this pathway is a central, linear information channel from the cell surface to the nucleus, it is a highly versatile one and itself part of a much more powerful Ras signalling network. Each cell can use the Ras network, together with the other oncogene-mediated signalling networks, to integrate and cross-regulate different membrane stimuli. A cell can thus receive a comprehensive picture of its environment, integrate this information into these powerful networks, and control both transient processes, such as changes in cell shape or motility, and less reversible processes such as differentiation, division and perhaps even programmed cell death.

S.E.E. & R.A.W.

MEK³⁸, which controls a third kinase, the mitogen-activated kinase termed MAP kinase^{16,39} (MAPK). This kinase and one yet further down in the cascade send signals into the cell nucleus by phosphorylating transcription factors that include products of oncogenes^{40,41}. It is this group of transcription factors that ultimately receives the signal and, in response, activates banks of genes. The chain of command is now almost complete, reaching from outside the cell through its surface membrane, through its cytoplasm, all the way to the cell nucleus.

The nomenclature is daunting, but the concept is clear. Proteins from an astonishingly wide range of research areas are all now arrayed in a single pathway. Further biochemical studies will no doubt identify both convergent and divergent arms. But, for the moment, those interested in the molecular mechanisms governing diverse processes including cell proliferation and differentiation can delight in schemes such as that depicted in the figure. Some details will require revision, but this overall outline appears to be rock solid.

The existence of this ubiquitous pathway means that those studying human cancer and vertebrate, fruitfly, worm or

even yeast development can benefit from each other's research. Many biologists working on obscure organisms have justified their research in terms of its benefits for cancer research. They can now take comfort from the fact that their labours indeed have direct implications for our understanding of how malignant transformation begins in human cells.

A circuit diagram has begun to emerge from the soup of the cell. The central elements are present even in simple, single-cell eukaryotic organisms such as yeast. Early in the evolution of multi-cell organisms, the peripheral elements involved in upstream signal acquisition were tacked on. Several downstream response elements and effectors were also hooked up. Once the main pathway illustrated in the figure was developed 700 or 800 million years ago, its machinery was exploited to govern various processes in myriad cell types.

This unification reveals — quite unexpectedly — that, in mammalian cells, extracellular signals associated with diverse biological processes call on this particular pathway to elicit many of their effects. Included here are the aforementioned growth-stimulatory factors (such as epidermal growth factor), insulin, anti-

gens bound by immunoglobulin cell-surface receptors, certain interleukins, and trophic factors (such as nerve growth factor). Evolution has created diversity and versatility by adapting this ancient pathway for a wide variety of purposes.

Stay tuned. The latest spate of work is only the beginning. In the coming years there will be hundreds and then thousands of papers describing how diverse proteins funnel their signals into this pathway. Other pathways, too, are being delineated, for example one involving direct targeting of transcription factors that are themselves tyrosine kinase substrates (see ref. 42). But the satisfaction to be had from the recent discoveries will endure — that of reducing extraordinarily complex phenomenology down to simple, apparently universally relevant, truths. □

Sean E. Egan is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK, and the Hospital for Sick Children Research Institute, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8. Robert A. Weinberg is at the Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA.

Preparing for the comet crash

Clark R. Chapman

THIS is the lull before the storm. Until late December this year, the planet Jupiter and a notorious nearby comet will remain temporarily hidden behind the Sun. Less than six months later, between 18 and 24 July 1994, at least 22 fragments of Comet Shoemaker–Levy 9 will crash into Jupiter — on the nightside, sad to say, so we will not have the full benefit of the attendant fireworks (see my News and Views article, *Nature* **363**, 492–493; 1993). Astronomers

Chicago), “a fireball as hot as the Sun will rise above the cloudtops.”

Smaller debris has spread far away: there are ‘wings’ that stretch beyond the visible train of fragments plus a fan of dust blown away from the nuclei by the Sun. Interactions between the dust and Jupiter’s magnetosphere, and atmospheric impacts of small fragments at the ends of the wings, could start two or three months before the main event.

and might last for weeks. And a quantity of Jupiter’s lower atmosphere mixed with the disintegrated comet — equal to the mass of each impactor — could be ejected into Jupiter’s stratosphere, contaminating it, changing its chemistry, and perhaps hazing over the southern hemisphere.

There is the possibility that the impacts will be big enough to excite waves that can probe Jupiter’s interior or elucidate the state of its atmosphere. As Tim Dowling (Massachusetts Institute of Technology) put it, the impacts will cause Jupiter to “ring like a bell”. If the vibrations and waves can be detected (for example if they induce clouds to condense and thus be-



The ill-fated ‘string of pearls’ comet. Left, Shoemaker–Levy 9 as seen by the ground-based Spacewatch Camera of the University of Arizona, on 30 March 1993 (J. V. Scotti); the field of view is about 600,000 miles across. Centre and right, close-ups of the nuclei taken by the Planetary Camera on the Hubble Space Telescope on 1 July (H. A. Weaver and T. E. Smith, NASA).

have digested four months of data obtained since the broken comet’s discovery in March, and reported results at a conference in Boulder last week*.

Perhaps the most spectacular result so far is the portrait taken on 1 July by the Hubble Space Telescope, where each comet fragment is a pinpoint of light enveloped in a spherical cloud of dust (see figure). But theoreticians at the meeting made bold predictions of what might happen during explosions the like of which have never been witnessed by human beings, and observatory directors unfolded plans for what may well become the largest campaign in the history of modern astronomy.

Most attention in Boulder focused on the ‘string of pearls’, the pieces of the original comet nucleus that was ripped apart by Jupiter’s tidal forces on 8 July 1992. Each of these sizeable fragments will flash brilliantly through Jupiter’s atmosphere at 60 kilometres per second and then explode far beneath its clouds with an energy probably exceeding a million megatonnes of TNT. “Within a minute after [each] explosion,” according to Kevin Zahnle (NASA Ames) and Mordecai-Mark MacLow (University of

The topic that undeniably dominated the marathon ‘comet crash’ scientific session on 18 October was the size of the comet fragments. (The size distribution of smaller debris and dust is almost wholly unconstrained by data or theory.) The progenitor comet is no longer thought to have been 20 km across nor are the fragments 10 km. Intensive searches of plates taken before the tidal break-up last year show nothing larger than about 10 km in Shoemaker–Levy’s orbit. But there was much debate in Boulder about whether the fragments might be far smaller than 1 km (David Jewitt, University of Hawaii), or as large as 3–4 km, as calculated from the Hubble Space Telescope images (Harold Weaver, Space Telescope Science Institute).

The size range translates into an uncertainty of a factor of a thousand in the explosive energy, making it difficult (as just one example) for astronomers to plan their exposure settings. If the fragments are as large as Weaver thinks, then a wealth of physical phenomena can be expected, including some that could be witnessed by amateur astronomers using modest backyard telescopes. There could be visible ‘light echoes’ of the impacts reflected from Jupiter’s satellite. Visible cloud spots formed by impacts would rotate into view a few hours after impact

come visible to cameras on the Hubble Space Telescope) insights might be gained that scientists never expected to be possible. Seismologists have long probed the Earth’s interior by detonating explosives to cause it to ‘ring’. This time, nature is providing us with a controlled explosive experiment to probe Jupiter.

On the other hand, astronomers have been burned by previous unfulfilled predictions of spectacular events, and some worry that small cometary fragments could fall into Jupiter without a trace. Last week most experts were optimistic. For one thing, the Galileo spacecraft has very sensitive detectors, which should easily detect jovian meteors much fainter than even the most pessimistic estimates — provided that the impacts are not just below the horizon. Second, there are circumstantial reasons for believing that the larger comet fragments must be at least a kilometre in size.

The strongest evidence comes from an unlikely place: the surface of Jupiter’s outer large moon, Callisto. Callisto bears scars which are almost certainly fossil records of earlier comet break-up events (P. Schenk and J. Melosh, *Nature* **365**, 731–733; 1993). Voyager photographs reveal a dozen crater chains on Callisto’s Jupiter-facing side. They look just like what would have resulted if Callisto had

*Annual Meeting of the Division for Planetary Sciences, American Astronomical Society, Boulder, Colorado, 18–22 October 1993.

Keeping out the rain

William M. Bement and Mark S. Mooseker

been in the way in July 1992 as Shoemaker-Levy 9, newly torn apart, headed away from Jupiter. From the tiny target area subtended by Callisto and the number of crater-chains formed during Callisto's 4,000 million years of exposure, Schenk and Melosh calculate that a comet must be torn apart by tidal forces in Jupiter's Roche zone about once a century. The co-discoverer of the comet, Eugene Shoemaker, who has just retired from the United States Geological Survey and joined the staff of Lowell Observatory (Flagstaff, Arizona), confirmed their estimate by an independent calculation based on numbers of comets known to be temporarily captured in Jupiter orbit: one break-up per century. As typical chain craters are about 15 km across, broken-comet fragments must approach 1 km in size.

During the next two months, many of the world's largest observatories will evaluate proposals to observe the impact next July. But the traditional mode of independent experiments by isolated, entrepreneurial astronomers is not appropriate for this once-in-a-millennium opportunity to witness a comet crash. Most observatories, including the Hubble Space Telescope (due to be repaired this December), will form teams to organize efficient observing programmes and to coordinate with observatories around the world; observations from every 30° of longitude would ensure that all 22 events are recorded. Communication networks must be arranged, instruments modified and command sequences radioed up to distant spacecraft.

When Jupiter and comet re-emerge from behind the Sun, the pace will become frantic. Exact positions of the comet fragments will be needed rapidly so that commands can be designed and transmitted to the two spacecraft best positioned to witness the impacts directly. Voyager 2 can look directly at the impact sites (so inconveniently placed on the far side from Earth) from its position in the far outer Solar System. Galileo, *en route* to Jupiter, can watch from the side as the comet fragments strike the giant planet's limb. Groundbased astronomers from the International Jupiter Watch will try to characterize Jupiter during the first months of 1994 so that they have a baseline against which to compare any atmospheric changes produced by the impacts.

And then everyone will wait to see what happens. None of the theoreticians in Boulder, I must say, seemed confident that their predictions would be confirmed. For never-before-seen events like the comet crash, I think we can confidently expect the unexpected. □

Clark R. Chapman is at the Planetary Science Institute, Tucson, Arizona 85705, USA.

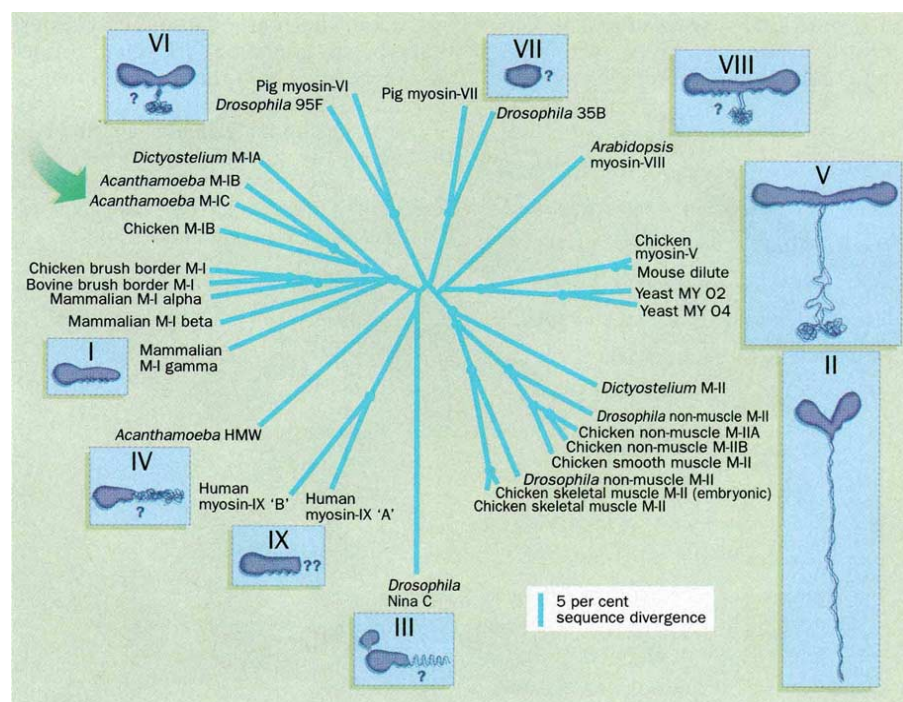
THERE is considerable movement in the field of molecular motors these days. The crystal structure of the myosin-II head has been solved¹, the biophysical properties of myosins^{2,3} and kinesins⁴⁻⁶ are being characterized in exquisitely fine (and often contradictory) detail, and new motors are being identified at an astonishing rate⁷. Yet this abundance of molecular information conceals a corresponding dearth of functional information, particularly in the case of 'unconventional' myosins, the nonfilamentous counterparts of the better-known 'conventional' myosins-II. Unconventional myosins are comprised of at least eight evolutionarily distinct classes^{7,8} (see figure), only two of which (myosins-III and -V) have representatives with clear functions based on mutational analysis. On page 841 of this issue⁹, however, a paper from the Korn and Pollard laboratories describes the functional characterization of *Acanthamoeba castellanii* myosin-IC, a member of the first- and largest-known class of unconventional myosins. (Indeed, this year is the twentieth anniversary of the discovery of *Acanthamoeba* myosin-I by Pollard and Korn¹⁰.)

The results of this study are especially gratifying in that, since their initial isolation from *A. castellanii*, myosins-I have been proteins in search of respect. Initially

disregarded by some as a proteolytic degradation product of the larger myosins-II, they finally gained ground in the 1980s through several independent lines of research. Specifically, it was found that not only are myosins-I diverse and widely expressed, but they also possess enzymatic and cellular properties that make them outstanding candidates as mediators of actin-based membrane movement within cells¹¹. The discovery that *Dictyostelium discoideum* without the myosin-II gene were still capable of cell locomotion¹² led to the logical conclusion that myosins-I probably power cell translocation.

So, by the end of the 1980s, myosins-I had satisfied most of the basic requirements for respectability: they were ubiquitous, numerous and probably key players in an essential cellular process. It was time to bring the power of homologous recombination to bear on them and demonstrate that they have a vital role in the life of the eukaryotic cell. Remarkably, however, knockouts of myosins-IB (refs 13,14) and -IA (ref. 15) in *D. discoideum* resulted in completely viable cells, and only relatively modest defects in cell motility.

Re-enter *A. castellanii*, and a combination of astute target selection and powerful technologies. Doberstein *et al.*⁹ have taken advantage of the observation that



Phylogenetic analysis of the myosin superfamily based on sequence comparison of the motor or head domain of known myosins. Question marks indicate either hypothetical or unknown structural features. Only a fraction of known myosins-I and -II are shown; the position of *Acanthamoeba* myosin-IC is indicated by the arrow. (Modified from ref. 8.)

A. castellanii myosin-IC is the only myosin localized to contractile vacuoles¹⁶, and of the technique of fluorescence activated cell sorting (FACS)-aided antibody loading. By syringe-loading *A. castellanii* with myosin-I antibodies in the presence of a fluorescent marker, they obtained large populations of cells in which the relative amount of loaded antibody could be quantified. Then, because the contractile vacuole is known to regulate water content in these soil amoebae, cells were subjected to osmotic shock — the laboratory equivalent of a torrential downpour. The results are beautifully clear: antibodies that inhibit myosin-IC phosphorylation and activity *in vitro* perturb vacuole contraction and cause eventual cell lysis in a manner directly correlated with both relative antibody concentration and osmotic shock. Moreover, as antibodies directed against myosin-IB have no effect above background, the cell lysis phenotype is specific to anti-myosin-IC antibodies.

Plainly, *A. castellanii* myosin-IC is essential to the cell. But what do these results say about unconventional myosins in general? Because the mechanism by which contractile vacuoles regulate cellular water content is not understood, interpretations of the lysis phenotype represent a microcosm of the various functions speculatively ascribed to unconventional myosins, including regulation of the actomyosin-membrane interactions, actin-based organelle transport and regulation of specific membrane proteins^{7,11,13}. For example, the most straightforward explanation is that myosin-IC simply drives actin-based vacuole contraction, thereby expelling the

water contained within into the extracellular medium. Myosin-IC could also transport elements of the spongione, an organelle system that links water transport from the cytoplasm to the contractile vacuole. A more exotic possibility, analogous to that proposed for myosin-I in adaptation to auditory stimuli in hair cell stereocilia¹⁷, is that myosin-IC directly regulates a membrane protein involved in water transport.

The study also bears on the question of myosin diversity. Not only are there many unconventional myosin classes, but several representatives of different classes are common within a given cell^{7,13,18}. Two general hypotheses have been presented to explain myosin diversity, particularly with respect to myosins-I. The 'functional redundancy' hypothesis suggests that myosins-I overlap in essential cellular functions, so that loss of a given myosin-I is compensated for by the others. The 'one myosin, one organelle' hypothesis suggests that each myosin carries out an organelle-specific function. The results of myosin-I knockouts in *D. discoideum* are consistent with the first hypothesis, whereas those of Doberstein *et al.*⁹ pro-

vide the first strong support for the second. The results of Doberstein *et al.* also suggest that subjecting *D. discoideum* myosin-I knockouts to specific challenges may reveal defects not detected under laboratory culture conditions which are a far cry from life in the dirt.

Finally, the *Acanthamoeba* system holds tremendous promise for solving a key problem in the study of unconventional myosins: identification of docking proteins which link myosins to membranes. Most unconventional myosins described to date are thought to have domains that specify phospholipid binding^{7,10}, and yet their patterns of membrane localization tend to be very specific¹⁶, meaning that information beyond simple phospholipid binding may dictate their distribution *in vivo*. Isolation of contractile vacuoles and analysis of myosin-IC binding proteins found in them should therefore provide welcome insight into unconventional myosin sorting. □

William M. Bement and Mark S. Mooseker are in the Department of Biology, Yale University, New Haven, Connecticut 06511, USA.

METEORITICS

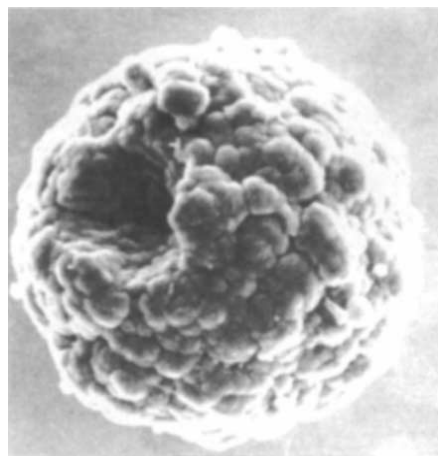
Ingrained evidence of origin

Ian Wright

THE Murchison meteorite, which fell in Victoria, Australia, in 1969, is one of a brood known as carbonaceous chondrites that typically contain 2–3 per cent carbon by weight, mostly in the form of organic materials¹. This carbon complex can be isolated by using acids such as HCl and HF to destroy the surrounding silicates and hydrated minerals. Within the carbon-rich residue it is possible to find tiny grains of diamond, silicon carbide and graphite, all

of pre-solar origin². These grains are survivors from the primaevial dust cloud that coalesced some 4.5 billion years ago from the solar nebula. On page 806 of this issue Amari *et al.*³ report chemical and isotopic measurements on graphite from the Murchison meteorite in an attempt to explain their origins.

To begin with, consider the enormity of the analytical task. First, the graphite grains in question constitute only 2 parts per million of the meteorite, and have to be isolated from a carbon-rich residue; it is rather like trying to find a particular straw in a haystack. Not surprisingly, then, it has taken nearly 20 years of refinement to arrive at the appropriate separation techniques⁴. Second, to discern the origins of the different sorts of graphite it is necessary to analyse individual grains. These are nominally about 1 micrometre in size, being composed of about 10¹¹ carbon atoms. An effective investigation requires not only the determination of their carbon stable isotopic compositions (¹²C/¹³C), but also of the isotope ratios of various trace constituents, such as nitrogen, oxygen and magnesium (this last to assess the abundance of once-radioactive ²⁶Al). The procedure must be repeated for many grains if populations are to be distinguished on the basis of statistically significant variations in composition.



Scanning electron micrograph of an interstellar graphite grain from the Murchison meteorite (about 11 µm across). (Courtesy of E. Zinner.)

1. Rayment, I. *et al.* *Science* **261**, 50–58 (1993).
2. Ishijima, A., Doi, T., Sakurada, K. & Yanagida, T. *Nature* **352**, 301–306 (1991).
3. Uyeda, T. Q. P., Warrick, H. M., Kron, S. J. & Spudich, J. A. *Nature* **352**, 307–311 (1991).
4. Kuo, S. C. & Sheetz, M. P. *Science* **260**, 232–234 (1993).
5. Hall, K., Cole, D. G., Yeh, Y., Scholey, J. M. & Baskin, R. J. *Nature* **363**, 457–459 (1993).
6. Svoboda, K., Schmidt, C. E., Schnapp, B. J. & Block, S. M. *Nature* **365**, 721–727 (1993).
7. Mooseker, M. S. *Curr. Biol.* **3**, 245–248 (1993).
8. Cheney, R. E., Riley, M. A. & Mooseker, M. S. *Cell Motil. Cytoskel.* **24**, 215–223 (1993).
9. Doberstein, S. K., Baines, I. C., Wiegand, G., Korn, E. D. & Pollard, T. D. *Nature* **365**, 841–844 (1993).
10. Pollard, T. D. & Korn, E. D. *J. Biol. Chem.* **248**, 4682–4690 (1973).
11. Pollard, T. D., Doberstein, S. K. & Zot, H. G. A. *Rev. Physiol.* **53**, 653–681 (1991).
12. DeLozanne, A. & Spudich, J. A. *Science* **236**, 1086–1091 (1987).
13. Jung, G. & Hammer, J. A. III. *J. Cell Biol.* **110**, 1955–1964 (1990).
14. Wessels, D., Murray, J., Jung, G., Hammer, J. A. & Soll, D. R. *Cell Motil. Cytoskel.* **20**, 301–315 (1991).
15. Titus, M. A., Wessels, D., Spudich, J. A. & Soll, D. *Molec. Biol. Cell* **4**, 233–246 (1993).
16. Baines, I. C., Brzeska, H. & Korn, E. D. *J. Cell Biol.* **119**, 1193–1203 (1992).
17. Gillespie, P. G., Wagner, M. C. & Hudspeth, A. J. *Neuron* (in the press).
18. Titus, M. A., Warrick, H. M. & Spudich, J. A. *Cell Regul.* **1**, 55–63 (1989).

Graphite grains come in a variety of morphologies, but only the rounded ones (see micrograph on facing page) appear to have large anomalies in their $^{12}\text{C}/^{13}\text{C}$ ratios. Amari and colleagues find that these range from 2 to 7,300 (a typical Solar System value is 89). The isotope ratios fall into several distinct groupings. The question facing experimentalists and astrophysicists alike is, how many sources of graphite does this represent?

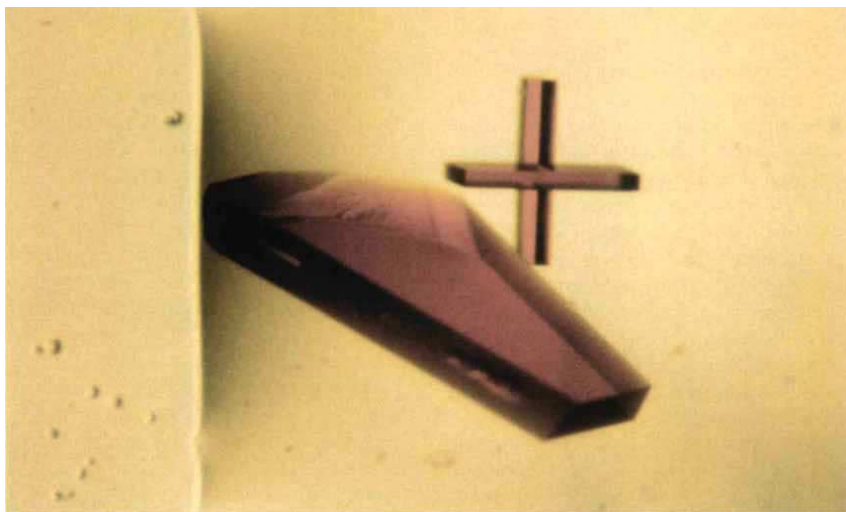
Similar questions are being asked for those other carbon-bearing interstellar dust grains, diamond and silicon carbide. In the case of diamond, for which the $^{12}\text{C}/^{13}\text{C}$ ratios vary only slightly from one meteorite to another⁵, and the size distribution reflects a low degree of processing⁶, there need be only a single source (two at most). However, because the grains are so small (typically 1 nanometre across), analyses are done on millions of crystals rather than individuals, so the measured isotope ratios might simply represent the average of many sources of pre-solar dust. For silicon carbide grains, which are sometimes large enough to be analysed individually, isotopic information has been variously interpreted as representing one source⁷, or up to 75 (ref. 8).

To settle the question, what is needed is the isolation of several grains that demonstrably come from a single source, and the newly discovered aggregates of several hundred silicon carbide grains⁹ may fit the bill. For various reasons it is unlikely that the aggregation took place in the solar nebula; it is more likely to have happened close to the source of the grains. The overall isotopic compositions of the clumps are close to the average of earlier individual analyses, so — if aggregation during the chemical treatments that produced the grains can be ruled out — it looks as if the range of isotopic values seen previously could have come from a single source.

So what of the variations in $^{12}\text{C}/^{13}\text{C}$ for graphite? Essentially, graphite, like silicon carbide, will condense from a stellar source where the carbon/oxygen ratio is greater than 1. Carbon stars at the asymptotic giant branch phase of evolution seem good candidates for the origin of silicon carbide. Curiously, though, the carbon isotopic compositions of graphite tend not to agree with those recorded in individual silicon carbide grains. Nor do they agree with those from carbon stars — puzzling, given that these are thought to contribute most of the carbonaceous dust to the interstellar medium. Why should the solar nebula have sampled an unrepresentative part of the interstellar medium?

For those graphite grains that have high $^{12}\text{C}/^{13}\text{C}$ ratios it is clear that nucleosynthesis via helium-burning is required. Amari and colleagues suggest two new possibilities: Wolf-Rayet stars, at a period in their

Halloween crystals



THESE macabre crystals formed over Halloween week-end last year. But sad to say neither supernatural forces, nor ingenious trick-or-treaters, have yet been implicated in their production. The cross (arms 0.5 mm in length) and coffin (complete with hinges) were a one-off, bizarre result of experiments involving crystallization of complexes of horse cytochrome *c* and its antibodies by vapour diffusion in hanging drops; normally the result is crystals 0.5 mm long resembling nothing so much as the everyday rock. The group concerned — a collaboration between researchers at the Universities of Minnesota and Illinois at Chicago, and Argonne National Laboratory — will test the reproducibility of the cross-and-coffin effect this Sunday. Ghostbusters are on standby. Those interested in the earthly aspects of the work can read about them in a paper by Ronald Jemmerson and colleagues, now in press with *Acta Crystallographica*. T.L.

evolution when the products of helium-burning rise to the surface; and helium-burning in the outer shells of massive stars, before they explode as type II supernovae. The accompanying isotopic evidence ($^{14}\text{N}/^{15}\text{N}$, $^{16}\text{O}/^{18}\text{O}$, ^{26}Al) does not, sad to say, enable a distinction to be drawn between these possibilities. It seems unlikely, though, that the large sizes of some of the graphite grains could be accounted for by supernovae.

It looks like graphite will turn out to have multiple sources, but for a complete understanding of those sources, we need still more isotopic measurements. In particular, noble gases seem to be essential. Krypton isotope measurements¹⁰ pin down the origin of some of the graphite grains to a pulsing star in the asymptotic giant branch, and neon from another subset of grains bears the signature of a nova explosion¹¹.

Those less interested in sources than in the nature and mixture of the dust grains that were drawn into the solar nebula will want to know the average isotopic compositions of meteoritic graphite. For them, the painstaking analysis of many hundreds of individual grains may not be the quickest way forward, and progressive combustion techniques of the type employed by Ash *et al.*¹² may be preferable. Ash and co-workers discovered a component of carbon in the Allende carbonaceous chondrite characterized by $^{12}\text{C}/^{13}\text{C}$

of 120. With hindsight, this isotopic value probably represents the average of a single population of meteoritic graphite. Yet a further population can be distinguished by the component known as C α , which has a $^{12}\text{C}/^{13}\text{C}$ ratio of 66 (ref. 13). These numbers can now be used alongside average $^{12}\text{C}/^{13}\text{C}$ values for diamond (92.5) and silicon carbide (37), in chemical models of Solar System formation. In the meantime, Amari and colleagues must continue their demanding task so as to satisfy the appetites of astrophysicists intent on describing the nucleosynthetic histories of Solar System materials. □

Ian Wright is in the Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

- Mullie, F. & Reisse, J. *Topics Curr. Chem.* **139**, 83–117 (1987).
- Ott, U. *Nature* **364**, 25–33 (1993).
- Amari, S., Hoppe, P., Zinner, E. & Lewis, R. S. *Nature* **365**, 806–809 (1993).
- Anders, E. in *Meteorites and the Early Solar System*, 927–955 (University Arizona Press, 1988).
- Russell, S. S., Arden, J. W. & Pillinger, C. T. *Science* **254**, 1188–1191 (1991).
- Lewis, R. S., Anders, E. & Draine, B. T. *Nature* **339**, 117–121 (1989).
- Stone, J., Hutcheon, I. D., Epstein, S. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **107**, 570–581 (1991).
- Alexander, C. M. O'D. *Geochim. cosmochim. Acta* **57**, 2869–2888 (1993).
- Russell, S. S. *et al.* *Meteoritics* **28**, 425–426 (1993).
- Lewis, R. S. & Amari, S. *Lunar planet. Sci.* **XXIII**, 775 (1992).
- Amari, S. *et al.* *Nature* **345**, 238–240 (1990).
- Ash, R. D. *et al.* *Nature* **336**, 228–230 (1988).
- Tang, M. *et al.* *Geochim. cosmochim. Acta* **52**, 1221–1234 (1988).

Pluto's strange orbit

S. J. Peale

THE planets in the Solar System are believed to have accumulated in a dusty gaseous disk, where internal dissipative processes caused the disk material to be flattened and relaxed into circular motion around the Sun. As a consequence, most have orbits that are near-circular and nearly coplanar. But Pluto is an oddity, as Bill McKinnon recently pointed out¹; not least among its peculiarities is its orbit, which is inclined by 17° to the plane of the ecliptic, and is so eccentric that it overlaps that of Neptune (see box).

How could such a bizarre orbit have come about as the planets formed in the plane of their dusty disk? Renu Malhotra², on page 819 of this issue, suggests that Pluto was born with the usual circular orbit in the plane of the disk. Then, she says, Neptune's orbit expanded as it scattered unaccreted planetesimals out of the region, and its decreasing orbital angular velocity approached 3/2 that of Pluto (for such a ratio of two small integers, the orbital angular velocities are said to be commensurate). Pluto was captured into the first of the orbital resonances described in the box, and it was subsequently pushed out ahead of Neptune, staying within the resonance while its eccentricity continued to grow as long as Neptune's orbit continued to expand.

The unique feature of this explanation is that it does not rely on a collision between Pluto and another planetesimal to push Pluto directly into the stable librating state. It readily produces the observed eccentricity of orbit and amplitude of libration, provided that Neptune's orbit expanded by about 5 or 6 astronomical units (1 AU is the mean distance from the Earth to the Sun). This migration is typical of that given by numerical models of the scattering of planetesimals by the four major planets during the last phases of planet formation³.

Orbital resonances abound in the Solar System, mostly among the satellites of the major planets, and capture of two bodies into this sort of resonance as their orbits approach one another through some dissipative process is well understood⁴. For satellites of the major planets, that dissipative process causing differential expansion of the orbits is tidal friction in the primary. Here Neptune's orbit approaches that of Pluto because some fraction of those planetesimals that Neptune scatters inwards are scattered out of the Solar System by Jupiter or Saturn, and the energy and angular momentum that Neptune gains by this process is more than the energy and angular momentum it loses by scattering

planetesimals outwards. As Neptune approaches commensurability with Pluto, capture is assured if Pluto's initial orbital eccentricity is less than 0.03, and is still possible for larger eccentricities.

Malhotra's explanation is not complete, as the inclination of 17° is not produced, nor is the resonance that keeps the apheion 90° away from the ascending node evident. But her numerical integration shows Pluto's inclination increasing, so there is a real possibility that some combination of initial conditions will yield all the properties of Pluto's orbit. Growth in the inclination might be due to another resonance associated with the 3/2 commensurability, but affecting the inclination rather than the eccentricity. Alternatively, it may be due to a 'secular' resonance, where the motion of the longitude of Pluto's ascending node is commensurate with one of the fundamental frequencies of the Solar System. This process of inclination growth within resonances has been used to explain the 4° inclination of

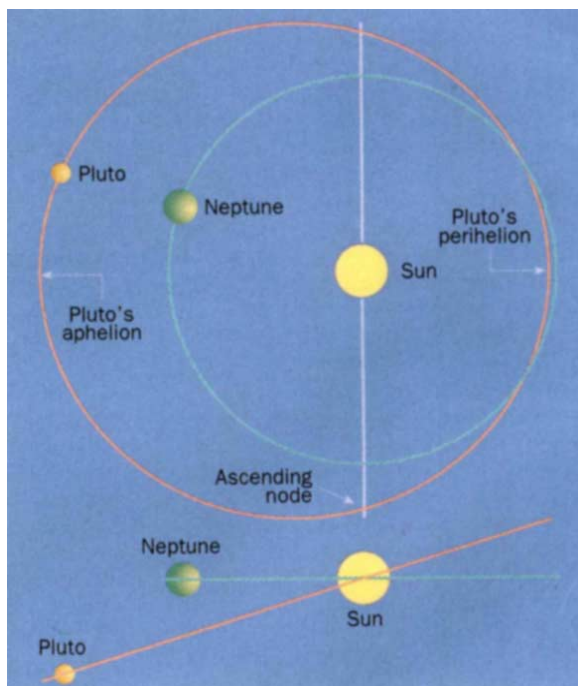
the orbit of Uranus's satellite Miranda in an otherwise planar system of satellites^{5,6}.

What of other explanations? The old theory that Pluto was a satellite of Neptune that escaped in a close encounter with Neptune's large satellite Triton⁷ is now thought to be unlikely^{8,9}. But Pluto could have been one of a swarm of similar planetesimals, some fraction of which were continuously scattered into and out of the 3/2 orbital resonance with Neptune by their mutual collisions. Pluto was simply left behind in its protective resonances after the others were scattered from the region by close approaches to Neptune.

This idea has gained support from calculations by Holman and Wisdom¹⁰ and by Levison and Stern¹¹. They independently find that if test particles are started in circular orbits nearly coplanar with Neptune's orbit with a range of semimajor axes near that corresponding with the 3/2 orbital resonance, the particles rapidly and chaotically evolve to eccentricities and inclinations like Pluto's. Some of these test particles are actually librating within the 3/2 resonance, but with such large amplitudes that they are chaotically unstable. A necessary collision between

Avoiding close encounters

PLUTO is usually the most distant planet from the Sun, but its orbit is so eccentric that it overlaps that of Neptune (in fact, Pluto is currently closer to the Sun). The two have avoided close approaches because Pluto's orbital period is almost exactly 3/2 that of Neptune and the planets are in conjunction when Pluto is near its aphelion, far away from Neptune's nearly circular orbit. The figure shows a typical conjunction between Neptune and Pluto. Rather than the longitude of the conjunction drifting away from Pluto's aphelion, it is pulled back by restoring accelerations from Neptune. This causes stable libration



(oscillation) of the conjunctions about Pluto's aphelion and the system is said to be in an orbital resonance.

A second orbital resonance keeps the aphelion roughly 90° away from Pluto's 'ascending node' (the point where Pluto's orbit rises through the plane of Neptune's orbit). So at conjunction, Pluto is close to its maximum distance below Neptune's orbital plane, and further still from Neptune.

Without at least the first of these resonances, Pluto would not now exist. The inevitable close approach to Neptune would either have produced a collision or thrown the smaller, lighter Pluto out of the Solar System.

S.J.P.

planetesimals within the resonance zone, which establishes a stable libration amplitude, has been shown to be possible¹⁰. That Pluto has suffered a major collision at some point may be implied by its present binary nature¹². The radius of its satellite, Charon, is about 620 km compared with Pluto's radius of 1,180 km (ref. 13) and the two are separated by only 19,405 km (ref. 14).

There are caveats for both models. At the time when there were abundant planetesimals available for collisions to nudge Pluto into a stable libration, the planets would not yet have migrated to their present positions as a result of their scattering of the planetesimals. The appropriate fundamental Solar System frequency might not then have been near that of Pluto's ascending-node motion, and Pluto could not have been trapped in the secular resonance that forced its inclination to high values. So if the growth in the orbital inclination seen in refs 10 and 11 is due to a secular resonance depending on the current configuration of the major planets, Pluto would have reached its current high inclination with an unstable libration after most of the other planetesimals in the region had been scattered away. The scarcity of planetesimals would have reduced the chance of the collision necessary to nudge Pluto into a stable libration in the short time (about a million years after Pluto became Neptune-crossing) before it made a close approach to Neptune.

The new model of Malhotra also needs a collision or some other scheme to create a binary planet. But collisions are almost inevitable during the time that Neptune is pushing Pluto's orbit outwards and they may ultimately be fatal to the model. If planetesimals are so easily captured into the 3/2 orbital resonance, then most planetesimals that encounter this resonance would be captured and have their eccentricities increased to large values. The orientations of these resonant orbits would be randomly distributed, and where their paths crossed, collisions would be likely to nudge the system out of resonance or be otherwise destructive. The pieces might be repeatedly scattered

out of the resonance before orbital parameters like those of Pluto's orbit are attained. So the survival of Pluto in resonance until the migrating Neptune has forced its eccentricity and perhaps its inclination to today's large values may be even less probable than the collisional emplacement of a chaotically inclined, eccentric Pluto directly into stable libration.

If initial conditions in Malhotra's model can be tweaked to produce the rest of the

characteristics of Pluto's orbit, especially the high inclination, it would be worth making additional extensive calculations to estimate probabilities for both models. A definitive choice between models may not result from this exercise, but perhaps a favourite will emerge for the history of this fascinating dynamical configuration. □

S. J. Peale is in the Department of Physics, University of California, Santa Barbara, California 93106, USA.

PALAEOANTHROPOLOGY

Rift on the record

Bernard Wood

GAPS in the palaeontological record are geographical as well as temporal. Almost all the fossil evidence for the early stages of human prehistory has come from sites in and around the eastern arm of the African Rift Valley or from southern Africa. The two regions are separated by nearly 3,000 kilometres but from similarities between the non-hominid mammalian faunas it seems that there must have been periods of contact. Schrenk and his colleagues in the Hominid Corridor Research Project (HCRP) have spent a decade surveying the areas between eastern and southern Africa, and on page 833 of this issue¹ they report on Plio-Pleistocene fossils they have recovered. Principal among those fossils is that of a jawbone, which can be attributed to the genus *Homo*.

Eastern and southern Africa are linked by the western arm of the African Rift Valley and a ribbon of lakes from Lake Albert in the north to Lake Malawi in the south mark its path. Until these latest discoveries, the Western Rift Valley had yielded only meagre and poorly preserved evidence of early hominids². But volcanic ashes in the Ugandan sites match ash layers in the Omo region in the Eastern Rift Valley, allowing the Omo sites to be fitted into the detailed chronological framework that has been developed for the latter region³.

The HCRP team made their discoveries in beds exposed along the western side of Lake Malawi. Vertebrate fossils were found at 131 localities, the greatest concentrations coming from around Karonga in the north and Uraha to the south (see map, page 833). These localities provide two 'windows' into the Plio-Pleistocene, namely Units 2 and 3, dating to more than 4 million years ago (Myr) and 3–1.5 Myr respectively. Even though the Malawi sites are geologically similar to those at Olduvai Gorge and the Omo Group in the eastern rift, it seems that they are beyond the range of the ash clouds produced by East African volcanoes (despite ash from

those volcanoes being recovered from deep-sea cores taken from the Gulf of Aden over 2,000 km east of Omo⁴).

It was from sediments of Unit 3A near Uraha that the project team recovered the early hominid mandible, UR 501. Fossils of large mammals, including pigs and antelopes, found in that subunit correlate with taxa at East African sites which date to between 3 and 2 Myr, the best fit being between 2.5 and 2.3 Myr. But the date of 2.4 Myr for UR 501 given by Schrenk *et al.* can only be taken as indicative until absolute ages can be obtained.

Early hominids from around 2.5 Myr are well known from both East and southern Africa. Two australopithecine species *Australopithecus afarensis* and *Paranthropus boisei* or *aethiopicus* are known in East Africa from this time, whereas there is just a single hominid species, *Australopithecus africanus*, recognized at the southern sites. What makes the Uraha mandible so intriguing is that it does not appear to belong with these australopithecine taxa but with a poorly known species, *Homo rudolfensis*. This is one of two hominid species that have been proposed to accommodate remains, principally from Olduvai Gorge and Koobi Fora, which had previously been loosely referred to as early *Homo* or *Homo habilis sensu lato*.

Taxonomic issues must detain us for a moment. I have proposed⁵ that the cranial evidence in the early *Homo* category should be subdivided into two species on the grounds of differences in the morphology of the brain case, face, jaws and teeth. One species, *Homo habilis sensu stricto*, is represented in the faunal records of both Olduvai and Koobi Fora by specimens such as OH 16 and 24 and KNM-ER 1813, whereas the other, *H. rudolfensis*, is only present at Koobi Fora. Fossils belonging to *H. rudolfensis* include the cranium KNM-ER 1470 and (crucial to the Malawi find) a well-preserved adult mandible, KNM-ER 1802.

The lack, in my judgement, of any

1. McKinnon, W. B. *Nature* **365**, 209 (1993).
2. Malhotra, R. *Nature* **365**, 819–821 (1993).
3. Fernandez, J. A. & Ip, W. H. *Icarus* **58**, 109–120 (1984).
4. Peale, S. J. in *Satellites* (eds Burns, J. A. K. & Matthews, M. S.) 159–223 (University Arizona Press, 1986).
5. Titterton, W. C. & Wisdom, J. *Icarus* **78**, 63–89 (1989).
6. Malhotra, R. & Dermott, S. F. *Icarus* **85**, 444–480 (1990).
7. Lyttleton, R. A. *Mon. Not. R. astr. Soc.* **97**, 108–115 (1936).
8. McKinnon, W. B. *Nature* **311**, 355–358 (1984).
9. Lin, D. C. N. *Mon. Not. R. astr. Soc.* **197**, 1081–1085 (1981).
10. Holman, M. J. & Wisdom, J. *Astr. J.* **105**, 1987–1999 (1993).
11. Levison, H. F. & Stern, S. A. (Paper presented at Pluto & Charon Conf., Flagstaff, Arizona, 6–9 July 1993).
12. McKinnon, W. B. *Astrophys. J.* **344**, L41–L44 (1989).
13. Millis, R. L. *et al.* *Icarus* **105** (in the press).
14. Null, G. W., Owen, W. M. & Synnott, S. P. *Astr. J.* **105**, 2319–2335 (1993).

evidence for *H. rudolfensis* at Olduvai is central to the taxonomic argument because the type specimen of *H. habilis* is a juvenile mandible and skull, OH 7, from Olduvai. So the species group that includes OH 7 must carry the name *H. habilis*. The suite of morphological differences that distinguish *H. rudolfensis* from *H. habilis sensu stricto* are functionally coherent and include a flatter, broader face and broader postcanine teeth with more complex crowns and roots and thicker enamel. Although *H. rudolfensis* crania are generally larger than those of *H. habilis sensu stricto*, the differences in overall size are a minor component of what is otherwise an impressive suite of morphological distinctions.

The same morphological evidence has been read differently by others and OH 7 judged to be distinct from the remaining non-australopithecine cranial evidence from Bed I and Lower Bed II of Olduvai Gorge^{6,7}. This results in an alternative taxonomy, with KNM-ER 1470 and 1802 joining OH 7 in *H. habilis sensu stricto* and KNM-ER 1813, together with OH 16 and 24, belonging to a second, unnamed, early *Homo* species. Schrenk *et al.* apparently accept the former taxonomic analysis at Olduvai and assign UR 501 to *H. rudolfensis* because of its resemblances to KNM-ER 1802 from Koobi Fora.

That said, the importance of UR 501 is threefold. First, its discovery supports the idea that at least one other hominid genus, in the form of *H. rudolfensis*, has paralleled *Paranthropus* in developing a morphology of its face, jaws and teeth which places emphasis on a capacity for chewing.

Second, together with the morphologically more tenuous evidence from Chemeron⁸, UR 501 suggests that *H. rudolfensis* arose as early as 2.5–2.4 Myr. This date coincides with what had been interpreted as an abrupt change in the global⁹, and particularly African¹⁰, climate from a warmer, moister regime to a colder and more arid one. But the move to increased aridity may have been a slower, incremental, process which began at around 3 Myr and continued through to 2 Myr (ref. 11). So attempts to correlate the evolution of African mammals with external climatic influences no longer have to demonstrate a strict 2.5 Myr synchronicity. Nonetheless, the Uraha evidence suggests that if increased emphasis on chewing was occurring in two hominid lineages, then the first appearances of the taxa involved, *P. aethiopicus* and *H. rudolfensis*, are both dated at around 2.5 Myr (ref. 12).

Third, the identification of UR 501 as *H. rudolfensis*, reinforces evidence from other large-mammal groups in Unit 3A that the Malawi faunas have stronger links with East Africa than with sites of similar age in southern Africa. From comparison of the mammalian faunas from the two

regions, Turner and myself¹³ concluded that between 4 Myr ago and now there have been both movements of fauna between the two regions as well as dispersions in common to them.

Schrenk *et al.* argue that the 'window' that Unit 3A provides onto faunal exchange implies that at around 2.4 Myr the predominant movement was from the southern to the eastern region. They infer this because, of the faunal elements from the Malawi site, more originated in the southern than in the eastern region. And they suggest that this ties in with the climate becoming more arid resulting in the Equatorward drift of the habitats which are then tracked by many mammals, particularly bovids. But first appearances of fossil taxa are a notoriously fickle tool to use to compare two regions with temporally discontinuous samples from different time ranges. The authors interpret the absence of *H. rudolfensis* from southern Africa as further evidence that evolutionary events were mainly occurring in the environments of East Africa where tropical forest and woodlands persisted, even during periods of relative aridity.

The special significance of the new discoveries is that they furnish evidence about the dynamics of the faunal interchanges between the two main African regions that have yielded early hominids. With analyses of deep-sea cores providing more precise records of past climates, evidence from sites such as Karonga and Uraha will lead to a better appreciation of the ways mammalian groups have responded to climatic fluctuations¹⁴. In particular we can hope for clarification of the relationships between the hominid fossil records of the two regions, and a better understanding of why hominid lineages responded in the ways they did to the climatic challenges of the Late Pliocene. □

Bernard Wood is in the Department of Human Anatomy and Cell Biology, University of Liverpool, Liverpool L69 3BX, UK.

- Schrenk, F., Bromage, T. G., Betzler, C. G. & Ring, U. *Nature* **365**, 833–836 (1993).
- Senut, B., Pickford, M., Ssemmanda, I., Elepu, D. & Obwona, P. C. *r. hebdomadaire Séanc. Acad. Sci., Paris* **305**, 819–822 (1991).
- Pickford, M., Senut, B., Poupeau, G., Brown, F. H. & Haileab, B. C. *r. hebdomadaire Séanc. Acad. Sci., Paris* **313**, 223–229 (1991).
- Sarna-Wojcicki, A. M., Meyer, C. E., Roth, P. H. & Brown, F. H. *Nature* **313**, 306–308 (1985).
- Wood, B. A. *Nature* **355**, 783–790 (1992).
- Stringer, C. B. in *Major Topics in Primate and Human Evolution* (eds Wood, B., Martin, L. & Andrews, P.) 266–294 (Cambridge University Press, 1986).
- Rightmire, G. P. *Am. J. phys. Anthropol.* **90**, 1–33 (1993).
- Hill, A., Ward, S., Deino, A., Curtis, G. & Drake, R. *Nature* **355**, 719–722 (1992).
- Shackleton, N. J. *Nature* **307**, 620–623 (1984).
- Vrba, E. S., Burckle, L. H., Denton, G. H. & Partridge, T. C. (eds) *S. Afr. J. Sci.* **81**, 224–275 (1985).
- Raymo, M. E., Hodell, D. & Jansen, E. *Paleoceanography* **7**, 645–672 (1992).
- Walker, A., Leakey, R. E., Harris, J. M. & Brown, F. H. *Nature* **322**, 517–522 (1986).
- Turner, A. & Wood, B. A. *J. hum. Evol.* **24**, 147 (1993).
- Vrba, E. S. *J. Mammal.* **73**, 1–28 (1992).

DAEDALUS

Safe passage

TELEPHONES, broadcast receivers, data-links and modems have no memory. If you do not catch and store a message at once, it is gone for good. So Daedalus proposes a cable which, on the 'bucket-brigade' principle of bubble memories and charge-coupled devices, combines transmission and storage. You feed data into it like marbles into a pipe, and they shuttle along to the other end. If the flow is interrupted, the data stay safely where they are until it resumes.

Daedalus's pipe will be one of those zeolite solid-state electrolytes through which ions can flow freely, and its marbles will be packets of ions. Such flowing ions should form a sort of one-dimensional micro-liquid; they must tend to spread out by mutual repulsion so the liquid will be compressible. DREADCO's chemists are seeking sets of contrasting ions which form mutually immiscible micro-liquids. Injected into the cable by pulsed electrodes, they will travel along its molecular channels as a stable sequence of packets, without mixing. All the usual packet-switching technology would work on this cable. At the far end, each ion-packet will be decoded by its characteristic deposition-voltage.

Ionic cable thus combines delay-line memory and digital transmission. Ions can be fired almost losslessly through a crystal lattice at many kilometres a second, so transmission could be almost electronically fast. But by broadening the cable at its ends, the ionic packets could be slowed to a relative amble. The resulting ion-reservoirs could hold many minutes or hours of compressed traffic.

Communications will be transformed. When the computer-net goes down, no information will vanish. It will merely sit in the wires, and start moving again when they are reconnected. Similarly, a message transmitted to an inactive telephone will no longer be lost, but will simply build up in the wire. Pick up the phone; the ions will resume their march and you will hear the stored message. Switch off the TV, and incoming programmes will build up in the download. You could let them out at your leisure, and at any rate you chose. In fast-forward mode the compressed ionic traffic would emerge at high speed; it could be throttled back for real-time viewing, or released in single-picture bursts for freeze-frame display. You could even push it back into the wire, and watch it emerge again. Video recorders, E-mail terminals, faxes, answering machines, all the complex data-traps of modern communications, will be replaced by the electronic equivalent of letters landing on the doormat, to be read at your leisure.

David Jones

German Bight storms analysed

SIR — In the debate concerning climate change due to increasing emissions of radiatively active gases into the atmosphere, many people are concerned about the possibility of an intensification of extratropical storms. Even though the International Panel on Climate Change¹ took a cautious stand in this matter be-

series, but homogeneous daily time series of wind for studying extreme events are rarely available because of changing procedures in observing, reporting or analysing the wind. The geostrophic wind (blowing parallel to isobars and representing the first-order approximation of the real wind) computed from pressure readings from a few stations

can be regarded as a proxy for the real wind. From this model, annual frequency distributions of daily wind can be obtained for periods of 100 and more years. Any trend in the wind statistics will be reflected in these geostrophic wind statistics.

We applied this approach to the German Bight, in the southeast part of the North Sea, where three stations have reported air pressure since 1876 (Fig. 1). The resulting time series of the 1, 10 and 50% percentiles of the annual distributions of geostrophic wind speeds (Fig. 2) stayed remarkably stationary, showing that this storm statistic has not changed in the German Bight in the past 100 years.

An alternative, and potentially more convincing, analysis would be to examine the historical weather maps and count the number of storms with a core pressure below certain thresholds. Such an analysis

has been done, revealing a substantial increase in the number of severe storms in the North Atlantic area². The problem with this approach is that it cannot distinguish whether changes are 'real' or whether they result from the ever-increasing quality of the operational analyses arising from more and better observations, more powerful diagnostic tools and other improvements in the monitoring of the state of the troposphere. A more detailed mapping of the pressure distribution, however, automatically yields deeper lows. On the other hand, the long time-series of pressure readings suffers from no substantial inhomogeneities because the measuring instruments (mercury barometers) and reading methods have remained unchanged for more than 100 years.

Heiner Schmidt

Seewetteramt Hamburg,
20359 Hamburg, Germany

Hans von Storch

Max Planck Institut für Meteorologie,
Bundesstraße 55,
20146 Hamburg, Germany

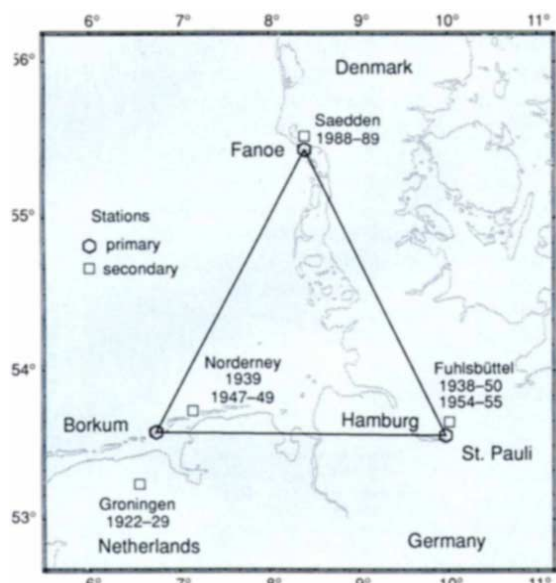


FIG. 1 Location of the meteorological stations from which pressure readings have been used to estimate geostrophic wind speed in the German Bight from 1876 to 1989.

cause of lack of evidence, a mixture of indirect evidence and misleading scientific statements has created substantial unease in the public.

Reports have been presented to the North Sea offshore oil industry about extreme waves, higher than ever previously observed. Two workshops 'Climate Trends and Future Offshore Design and Operation Criteria' in Reykjavik (29-30 March 1993) and Bergen (30 November-1 December 1992) were held in which the Norwegian Weather Service brought together people from the oil industry, certifying agencies and scientists to discuss whether the wave and storm climate is in fact worsening. No definitive answers emerged — it is not possible to tell whether the extremes had become worse or if reporting systems have improved. The insurance industry has organized meetings with scientists because of the increased number of claims for storm-related damage. Newspapers in northern Europe were full of speculation about the enhanced threat of extratropical storms in the early part of 1993.

To ascertain whether the storminess really had worsened, a systematic analysis was begun at the Seewetteramt Hamburg. Identification of a trend in climatic data requires long and homogeneous time

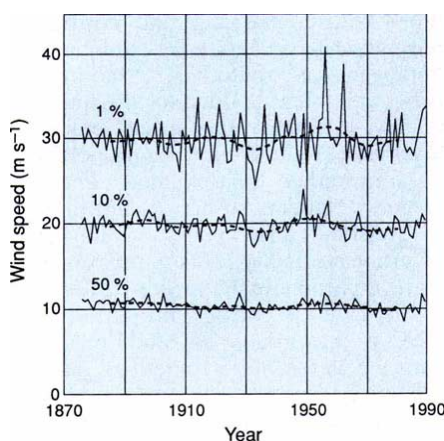


FIG. 2 Time series of the 1, 10 and 50% percentiles derived from annual distributions of daily geostrophic wind speeds in the German Bight. (The p 'percentile' of a distribution X is the number x , such that the probability of observing $X > x$ is p . Thus, the 1% percentile, derived from an annual ensemble of 365 samples, represents the minimum of the 3-4 largest observations.) Solid lines, annual average; broken lines, 30-year low pass.

Generation time and evolution

SIR — May in *News and Views*¹ examined the evolution of resistance to transgenic plants expressing genes for insecticidal proteins derived from the bacterium *Bacillus thuringiensis*. Based on a heuristic model², he concluded that the time required for an insect population to become resistant depends directly on the generation time of the insect but only weakly on other parameters (selection strength, initial resistance allele frequency and resistance allele dominance).

The influence of generation time on the rate of resistance evolution hinges on the relationship between generation time and the selection intensity per generation (s). May's analysis assumes that s is constant across species with different generation times. For insects exposed to a relatively constant dose of toxin expressed by a transgenic crop plant, this condition could be satisfied if only one developmental stage of the insect (for example, neonate larvae) is susceptible to the toxin. For insects exposed to conventional insecticides, this condition could be satisfied if the decision to apply insecticides is triggered by the density of the target pest population, and if the realized population growth rate is strongly influenced by generation time.^{3,4}

But it is also possible that s will be a

function of generation time. Under many field conditions, species with long generation times experience more intense pesticidal selection per generation than species with short generation times. If toxins cause a constant mortality rate per day, then the time for resistance to appear is independent of generation time, and instead, selection intensity becomes a primary determinant of the rate of pest adaptation⁴. Detailed numerical simulations show that generation time can interact with ecological and genetic factors to produce a variety of influences on the rate of resistance evolution⁵.

Species with long generation times may in some cases experience such intense selection from insecticides that population densities decline catastrophically and local extinctions occur⁶. In this case, the evolution of resistance in long-lived species may be retarded beyond that expected from population genetics models in which infinite population sizes are assumed⁷.

The extensive documentation of resistance evolution in arthropod pests provides an opportunity to test these different theoretical predictions. Our analysis of 586 North American arthropod pests of agriculture revealed no direct relationship between generation time and resistance

evolution⁴. The tendency to evolve resistance to synthetic chemical insecticides as of 1980 was not linearly related to generation time; species with intermediate generation times (between four and ten generations per year) evolved resistance more readily than species with shorter or longer generation times⁴, but generation time explained only a very small fraction (5.7 per cent) of the total variation in resistance evolution.

Here we report that an enlarged and updated analysis of resistance evolution in North American arthropod pests confirms the result that resistance evolution does not depend directly upon generation time. The ability of each species ($n = 607$) to evolve resistance was quantified as the number of insecticide classes to which resistance had been reported up to 1989 (ref. 8). The simplest logistic regression model suggests that the number of generations per year makes no contribution to the ability to evolve resistance ($\chi^2 = 3.2$, $P = 0.08$). A more complete model that incorporates the influences of pest severity and pest feeding mode (chewing, sucking on phloem and xylem contents, or sucking on cell contents) identifies a significant curvilinear effect for generations per year; species with intermediate numbers of generations per year displayed the greatest ability to evolve resistance (see figure).

The relationship between the generation time of pests and their evolution of resistance to toxins in transgenic plants cannot be fully elucidated until such plants are widely deployed. Selection exerted by

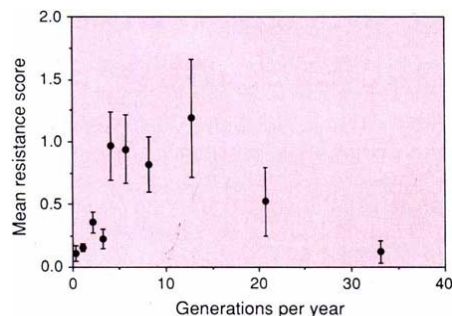
toxins produced in transgenic plants may differ from selection exerted by conventional insecticides. Nonetheless, existing evidence from conventional insecticides suggests that the role of generation time may be limited, whereas factors directly linked to the intensity of selection, such as the concentration and temporal and spatial distribution of toxins, are likely to be critical. This result is important because it shifts the focus of attention from a biological trait that one generally cannot control (generation time) to parameters that will be shaped directly by the manner in which insecticidal transgenic plants are deployed.

Jay A. Rosenheim

Department of Entomology,
University of California,
Davis, California 95616, USA

Bruce E. Tabashnik

Department of Entomology,
University of Hawaii,
Honolulu,
Hawaii 96822, USA



Resistance score (mean $\pm$ s.e.; defined as the number of insecticide classes to which resistance has been documented in the field) for arthropod pests of agriculture in North America as a function of the number of generations per year completed by the pest. Data were analysed using a polychotomous stepwise logistic regression model; both the linear term (generations per year; $\chi^2 = 4.0$, $P = 0.04$; coefficient = 0.31 ± 0.06) and the quadratic term ($[\text{generations per year}]^2$; $\chi^2 = 20.5$, $P < 0.0001$; coefficient = -0.0098 ± 0.0024) are significant. To facilitate data presentation, 10 intervals of generations per year are defined by the following cutoff points: ≤ 0.5 ($n = 27$); 1.5 (243); 2.5 (117); 3.5 (83); 4.5 (29); 6.5 (33); 10.5 (22); 15.5 (16); 25.5 (21); and > 25.5 (16). Values of generations per year were obtained from the literature. When estimates for the species in question were unavailable, average values for congeneric species were used ($n = 57$); for species in the families Aphididae and Tetranychidae, average values for confamilial species were used when congeners were unavailable ($n = 20$).

Thermodynamics at boundaries

SIR — After showing that entropy is proportional to mass or volume for homogeneous materials, Maddox in *News and Views*¹ addresses the question of how it is that, in a special case, the black hole, it turns out to be proportional to area rather than volume. The discrepancy appears to him quaint enough to welcome an explanation in which quantum mechanics, plus complex mathematical paraphernalia, are involved. But is the case of entropy proportional to area as uncommon as Maddox apparently believes, or something almost trivially common?

Take the case of the reversible association of two globular protein subunits, a subject that has occupied my attention for the past 10 years (ref. 2 and references cited therein). These associations are entropy driven, as uniformly found since Lauffer *et al.*³ discovered the first case in 1958. The only contribution to the entropy change on association comes from replacement of the bonds at the boundary between protein and water by bonds of protein with protein. Assuming equal

numbers of bonds of a typical energy for each reacting surface, the total entropy change on association is proportional to the number of bonds and therefore to the area of subunit interaction.

From such considerations it is easy to see that the same will be true of all cases of isothermal chemical reaction of polyatomic molecules that involve the bonds of the atoms at the interface and leave others unchanged: energy and entropy changes will be proportional to the interactive area, not to the volume of the reactive entities. Thus thermodynamics at the boundary of a black hole seem in this respect no different, Clausius and Gibbs be blessed, than at any other boundary.

Gregorio Weber

Department of Biochemistry,
University of Illinois,
Urbana,
Illinois 61801, USA

1. Maddox, J. *Nature* **365**, 103 (1993).
2. Weber, G. *J. phys. Chem.* **97**, 7108–7115 (1993).
3. Lauffer, M. A., Ansevin, A. T., Cartwright, A. T. & Brinton, C. C. *Nature* **181**, 1338–1339 (1958).

Blueprint for survival

Timothy O'Riordan

Compass and Gyroscope: Integrating Science and Politics for the Environment. By Kai N. Lee. Island: 1993. Pp. 243. \$24.95, £18.95. (European distribution, Earthscan.)

THE compass in the title is the guiding hand of integrative science in the pursuit of good environmental practice. The gyroscope is the sheet anchor of democracy and enlightened public support. Kai Lee was a senior analyst in the Columbia Basin Program, nowadays a major exercise in river basin management, the enhancement of salmonoid stocks and the generation of water-fed economic development in the Pacific Northwest of the United States. He uses his experience of negotiating deals in water use and environmental protection with disparate and, at times, suspicious interest groups as an illuminator for this thesis on the changing role of science in the uncertain search for the holy grail of sustainability.

He regards sustainable development as an ideal, a social good that will never be fully attained, but which should set markers and navigational aids in the long and sometimes troublesome voyage towards greater environmental permanence in human habitation of this beleaguered planet. Sustainability is just like democracy, freedom and justice — desirable means and idealized ends to which we should all strive for our own good. Lee also believes that these great concepts are interconnected. The sustainability revolution can only proceed alongside greater justice and democratic freedom for all species on Earth, human and nonhuman. An unfair and oppressed world cannot possibly become sustainable. Whereas wars may kill at worst millions, planetary instability could result in billions of lost lives and much in the way of species extinctions.

Lee warms to this theme by advancing the notion that science and democracy, in the wrong hands and with colonialist values, could well lead to a new Dark Age, "a miserable sustenance won with what skills survive, tainted by the legends of lost empires". The metaphorical end of the voyage of Columbus came, not with the discovery of the Pacific or the limits of the frontier, but with the vision of a blue, green and white orb floating in the blackness of inhospitable space, the one and only home for a species whose destiny lies in its "imperfect and limited means".

The metaphor of adaptation to this fresh reality, adopted by Lee, is that of a voyage of social learning through the application of "civic science" (the compass) and negotiated participation (the gyroscope). The navigational aid is a series of devices through which collective adaptation aimed at the common good

takes place. Rationality gives way to a series of biased judgements that are influenced by power, position, experience, expectation and political culture. Action takes place through "principals" or actors who lead, and "agents" or supporters who collaborate. Each is locked into the other by rules of compliance and mutual gain; at least, this should be the case if science is 'participatory' and the decision structure is open and accountable.

Lee coins the phrase 'civic science' to emphasize the important shifting role of science in an age of social adjustment towards sustainable survival. He regards science as having to be both credible and legitimate. Credibility stems from rigorous scientific methodology, set in a culture of learning and negotiation where proof of cause and effect is not always available, and where different possibilities for taking decisions require social judgement and not simply the views of experts. Heterogeneity of process at all levels of natural and social interaction means that even well-established scientific techniques such

as averaging sampled data, or the use of indicator species or circumstances as measures of tolerance or resilience, may no longer be acceptable. Science is still vitally needed; there is no gainsaying that. But the challenge is to devise ways of coping with uncertainty, namely probability of outcome, and indeterminacy, namely guesswork. Indeterminacy arises because the systems studied may operate chaotically and so are not amenable to conventional probabilistic analysis.

Legitimacy of science applies to the devices through which scientific exploration and prognoses allow otherwise unwilling parties to see a common advantage in realigning their biases in favour of a wider good. Civic science therefore acts as the lubricant for reformulated visions and relationships, working in conjunction with lively and educational democratic procedures. The latter could be anything from round tables of interested parties, focus groups of contending antagonists and negotiated learning through computer-aided images of future outcomes of certain courses of action, to science-based assessment procedures in any governing institution — local, regional, national or international.

Civic science has to be very careful about distributional outcomes of suggested change in economy and society. This is always the great determinant of the sustainable transition. Inevitably there



Assemblage of the Ordovician trilobite *Selenopeltis*, extracted from the Anti-Atlas mountain range in southwest Morocco and shown here at about a quarter of their natural size. The picture is taken from the jacket of the second edition of *Trilobites* by Riccardo Levi-Setti, which is packed full of photographs of these remarkable extinct marine arthropods. University of Chicago Press, \$45, £35.95.

will be a different group of gainers and losers from that in any pathway that is overtly unsustainable. Environmental taxation will not be popular unless the revenue is directed at those who are made worse off as a result, and at ensuring that the cause of the damage is altered by market signals and strategic subsidy. Such measures will not gain legitimacy in the absence of civic science. The transition to sustainability will be a pro-active and participatory experience of honest trial and error. There can be no blueprint nor any predetermined track. Hence the need for both a compass and a gyroscope.

This is a beautifully written book full of thoughtful insights into environmental policy-making by a person who is clearly sensitive to the political pragmatics of resource management and interest-group bargaining. The text is peppered with examples from the Columbia phenomenon — notably the argument over fish hatcheries, fish ladders, habitat protection for vulnerable species, pricing mechanism to alter the use and amount of water pumped into agriculture, industry, leisure and amenity, and the ever pres-

ent problem of changing biases from well-established prejudices concerning water-use rights and wealth-generating expectations.

Lee captures extremely well the subtle but exciting new roles for science and scientists in a post-Rio world. It remains to be seen how far his call for a more civic and interdisciplinary science will be heeded by those who allocate research budgets, fund training programmes and claim expertise and specialized knowledge. There is a real debate going on in all these areas now that contract-led science is becoming more fashionable and basic research is being starved of money. Civic science must not be the end-point of science: but it does offer a genuinely exciting role for a reformulated interpretation of science in an agonized world that is stumbling its way towards environmental survival. This book is for scientists at large, not just resource managers and technical environmentalists. □

Timothy O'Riordan is in the School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Social life from the bottom up

W. C. McGrew

Primate Behaviour: Information, Social Knowledge, and the Evolution of Culture.

By Duane Quiatt and Vernon Reynolds. Cambridge University Press: 1993. Pp. 322. £42.50, \$74.95.

HERE is at least the fourth book of this title, but its distinguishing features are revealed in the subtitle. Unlike the other three, which were empirically based or topically organized anthologies, this is an ambitious synthesis by two veterans of primatology. Duane Quiatt and Vernon Reynolds, respectively a US and a British biological anthropologist, attempt to weave together all levels of social life of nonhuman primates, and make a critical comparison of their findings with human primates. The theoretical perspective is broad, taking in both evolutionary and sociocultural anthropological perspectives, but artificial intelligence also makes a contribution. It is a 'high risk/high payoff' undertaking, to use the jargon of optimality theory.

Quiatt and Reynolds start with the premise that information transmission between individuals is the key to understanding social relations. Social cognition allows social knowledge, which underpins relationships, which in turn combine to determine social structure. (The authors rely heavily on the hierarchical view of social life espoused by Robert Hinde.) They carefully build a case, starting from

individual acts in their context, showing how tactics and strategies result in net benefits at the individual level of natural selection. At the same time, underlying perceptual and cognitive processes are manifest in social cognition (or 'mind-reading' as it is sometimes called). Intentional communication or miscommunication shapes relationships, as each individual primate competes or cooperates with her or his fellows.

The final third of the book focuses on human primates, taking the analysis to the highest level of social organization: institutions. Here the emphasis changes from continuities across the order Primates to discontinuities that define human uniqueness. Key parts are assigned to language, marriage, lineage, wealth and family. The conventionalizing of these roles allows the evolutionary transformation that culminates in human culture.

The attempt overall is gallant, but problems arise. Despite offering cogent definitions in principle of such phenomena as social knowledge, pedagogy and institution, these are not framed in empirically testable, operational terms. Thus, we are not told, for example, how to distinguish rule-governed from convention-governed behaviour in the real world, although the authors present this as a contrast between nonhuman and human beings. It may be that only humans assume true roles, but working primatologists might argue that this is very much a matter of definition.

This is compounded by the selective focus of attention among primate taxa, which reports findings almost entirely from Old World monkeys and from chimpanzees. Prosimians and New World monkeys are virtually ignored. Although this might have been justified until recently on the grounds of lack of data, that position is now outdated. To argue that fatherhood as a role is somehow essentially human ignores a mass of work from marmosets and tamarins showing sophisticated familial division of labour in childrearing.

The book's arguments are not well served by a somewhat pedantic writing style, too full of fancy words such as "evidencing", "processual", "solidary" and "juridicial". More curious are terms that seem to be idiosyncratic variations on standard ones: "provisionization" (provisioning), "pairedness" (pair-bonding), "degree of relationship" (relatedness), "matrikin" (maternal kin) and so on. This too often combines with obtuseness to produce a demanding test for the reader, as in: "It may help to address such questions if we make a distinction between the domain of social knowledge . . . and the social domain(s) of knowledge, or knowledge in the social domain".

The ideas in the book are notable for their skilful interweaving, rather than their novelty, with heavy debts to Nicholas Humphrey, Robert Trivers, Dorothy Cheney and Robert Seyfarth, and Richard Byrne and Andrew Whiten. More disturbing are the absences: Barbara King's work in the same vein is not mentioned, nor are David Premack and Guy Woodruff's trail-blazing studies on 'theory of mind' in chimpanzees.

The overall result is a useful synthesis of the current state of cognitive primatology, as seen through experienced anthropological eyes, that presents a challenge to readers seeking to make sense of this topical area of research. □

W. C. McGrew is in the Departments of Sociology and Anthropology, of Psychology and of Zoology, Miami University, Oxford, Ohio 45056, USA.

Biology in paperback

Foundations of Cognitive Science edited by Michael I. Posner. MIT Press, \$35, £24.95. For a review see *Nature* **351**, 281 (1991).

The Doctrine of DNA : Biology as Ideology by R. C. Lewontin. Penguin, £5.99. See *Nature* **363**, 27 (1993).

Life History: An Introduction to Taphonomy and Paleoecology by Pat Shipman. Harvard University Press, \$23.95, £15.95. See *Nature* **214**, 598 (1981).

Random Walks in Biology by Howard Berg. New edition. Princeton University Press, \$12.95, £10.95. See *Nature* **310**, 258 (1984).

Natural hazards

Peter W. Lipman

Earthquakes and Volcanic Eruptions: A Handbook on Risk Assessment. By Herbert Tiedemann. *Swiss Reinsurance Company, CH-8022 Zürich, Switzerland: 1992. Pp. 951. \$180.*

SCIENTISTS tend to be slow to communicate the results of their research in practically useful ways. Their desire for additional knowledge competes with the time, money and effort needed to develop applications for the benefit of society, and organizational and cultural barriers come between researchers and their potential customers. The considerable progress made during the past 50 years in our understanding of natural disasters is only gradually being exploited by governments in land-use regulation, by industry in investment decisions and by the general public in assessing personal risk. Recent catastrophes, such as the 1980 eruption of Mount St Helens, and the earthquakes in 1985 in Mexico City and in 1989 in Loma Prieta, California, the 1992 tropical storms in Florida and Hawaii and the 1993 floods in the Midwest of the United States, have heightened public awareness of natural hazards and vividly illustrated the fragility of a complex industrial society. One particular weakness is the limited capability of insurance systems to deal with such disasters, much less with the larger ones predicted on the basis of geological and meteorological data.

In this extraordinary volume on earthquakes and volcanoes — the most severe geological hazards that threaten human life and property — Herbert Tiedemann provides a major contribution towards bridging the information gap between the Earth-science community and those who need more economically effective risk assessments. Working with the Swiss Reinsurance Company, Tiedemann has generated a broad summary of knowledge about earthquakes and volcanic eruptions, principles that influence the resulting damage to society, and guidelines for assessing risks, intended for use by the insurance industry but broadly applicable to government and industrial policies and individual lifestyle decisions.

At 3.6 kg and measuring 22 × 27 cm, this tome is hardly a mere handbook. Rather, it is an impressive summary and synthesis that is handsomely produced, combining coated paper, attractive founts and many excellent illustrations (674 figures, most in colour). The index alone consists of 91 pages. Detailed data for risk assessment are compiled in 27 appendices, two of which are separate, colour global maps of historical earth-

quakes and volcanic eruptions (scale 1:23,000,000) and a paperback catalogue summarizing 697 individual events. One complaint: the binding is inadequate for a book of this size; the cover on my copy began to come away before this review was completed.

Tiedemann has read voraciously and on the whole thoughtfully; most of his summaries are thorough and up to date without being too technical. In places, however, basic principles are inadequately defined for the lay reader, who may be overwhelmed by the rambling discussion of details (on, for example, earthquake

largest known today”.

More convincingly, Tiedemann compiles historical data on large earthquakes (magnitudes greater than five) during the past century to conclude that the past 25 years have been seismically comparatively inactive. These data, in conjunction with the growth in global population, capital investment and use of high-risk land, suggest to Tiedemann that the cumulative exposure to insurance losses from earthquakes “is much larger than anticipated because of the gross underestimation of the vulnerability of elements at risk”. Indeed, the excellent photographs of

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Street incredibility — landslide in Anchorage caused by the Alaskan earthquake of 1964.

magnitude and intensity). Quite a few factual errors, even in figure captions, suggest that the volume was inadequately checked for technical accuracy. The lack of source citations, while in keeping with nontechnicality, is frustrating when the discussion involves alternative interpretations of primary data. At times, Tiedemann's enthusiasm for a topic leads to over-speculative conclusions. For example, in discussing large explosive eruptions of ash flows that lead to caldera collapse — probably the most catastrophic events to affect the Earth other than rare extraterrestrial impacts — Tiedemann confusingly combines size data for individual well-documented calderas, composite dimensions of multiple caldera collapses associated with recurrent eruptions (in, for instance, the Toba complex in Sumatra) and novel inference of even larger volcanic collapse structures based entirely on topographical features in Madagascar (where young ash-flow deposits are unknown). A resulting graph (Fig. 223; not based solely on data from Table 32 as claimed in the caption) is used to draw the dubious conclusion that “eruptions are possible which generate larger calderas (diameters in excess of 200 km) than the

structural damage from earthquakes, accompanied by captions discussing the diverse causes of failure, should be sobering to any building owner, insurance carrier or policy-maker.

I am ill-qualified to judge the sections on insurance applications, but such detailed use of Earth-science data for commercial risk assessment is to my knowledge novel and constitutes impressive progress beyond the arbitrary and socially damaging ‘red-lining’ approach taken by some insurers following recent natural disasters in North America. I know of no other comparable volume that so comprehensively combines a thorough summary of fundamental scientific results with direct applications to large-scale issues of public concern. Certainly, this volume should prove to be useful and interesting not only to insurers but also to planners, architects, engineers and those interested in risk management and optimization — in short, to anyone who wants a broad survey of current ideas about earthquakes and volcanoes. □

Peter W. Lipman is in the Volcano Hazards Program, US Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA.

Soviet revolutions

Harold Zirin

Alexander A. Friedmann: The Man Who Made the Universe Expand. By E. A. Tropp, V. Ya. Frenkel and A. D. Chernin. Cambridge University Press: 1993. Pp. 267. £30, \$49.95.

ALTHOUGH 'Friedmannian models' are widely used in cosmology, few people have any idea of the remarkable discoverer of these equations. Alexander Friedmann (1888–1925), the originator of the expanding Universe, was a balloonist, a highly decorated military aviator, an outstanding meteorologist and mathematician, and an administrator of some repute. This biography presents a fascinating picture of an active intellect and a turbulent life. Written for his centenary near the end of the Soviet era, the account is still somewhat Soviet, with the translation by Alexander Dron and Michael Burov following the original faithfully, for better or worse. But the story of this remarkable man is absorbing.

Friedmann came from a family of musicians. His grandfather converted from Judaism to the Russian Orthodox Church under the rigours of military service around 1830. His mother left his father when our hero was 12 months old, and as a result was sentenced to a life of celibacy by the church authorities. His father obtained a divorce in 1897 to marry a ballerina, and Friedmann was brought up by his grandfather. (These scandals are omitted from this book but noted in *A. A. Friedmann: Centenary Volume* edited by M. A. Markov, V. A. Berezin and U. F. Mukhanov (World Scientific, 1988).)

As a student at a gymnasium in St Petersburg, Friedmann excelled. The authors follow his education in exhaustive detail, even with examples of the complex matriculation mathematics problems he had to solve. He entered the University of St Petersburg as a student of academician V. A. Steklov. Despite the turmoil in the aftermath of the 1905 revolution, the 19-year-old Friedmann pursued the study of Bernoulli numbers, prime numbers, congruences and the theory of elementary functions.

Friedmann graduated in 1910 with a gold medal for a study of "second-degree indeterminate equations". He became a close friend of the Dutch physicist Paul Ehrenfest, who, with his Russian wife, then lived in St Petersburg. In 1913, Friedmann joined the aerological observatory in Pavlovsk, part of the Main Physical (later Geophysical) Observatory. This was the beginning of his long career in dynamic meteorology. Although the authors laud him as a representative of

'Soviet science', these achievements predated the Revolution; he was clearly a product of the old regime. Another annoying feature of the book, possibly due to the collaboration of three authors, is a tendency to lurch from one place or activity to another without explanation; in 1913 (p. 63) Friedmann suddenly turns up in Leipzig studying with V. F. K. Bjerknes without any explanation of how he got there.

Friedmann spent most of the First World War in military aviation, especially fighting the Austrians on the Galician front. His role was to make aerological observations, carry out reconnaissance and develop a practical theory of bomb trajectories, which he claims to have introduced successfully by bombing the arsenal of the fortress of Przemyśl. Steklov wrote to tell him that he should quit flying and concentrate on calculations, but he was eager to test his theories. The meteorologist H. Ficker, who fought on the Austrian side, recorded that the bomb dropped by Friedmann was the only Russian one to hit its target in Przemyśl. Friedmann was awarded the George Cross, the highest Russian military decoration. Despatched to Kiev as a teacher of meteorology to aviators, he there produced *A Synopsis of Lectures on Aeronautics*, a short book intended for use in training pilots. He then supervised the construction of a plant for manufacturing aviation instruments in Moscow.

The description of his adventures in the Russian Revolution and subsequent Civil War is confusing, perhaps deliberately so. Friedmann took part in the establishment of a new university in Perm, far to the east of Moscow. The city passed back and forth between the Bolsheviks and the Whites; the supporters of each side withdrew and re-entered accordingly. Friedmann was sympathetic to the Bolsheviks, but returned without trouble to Perm before Kolchak left. It is hard to understand what happened.

After the Revolution, Friedmann returned to the Main Geophysical Observatory and wrote his dissertation on *The Hydrodynamics of a Compressible Fluid*. This important work had great influence on dynamic meteorology in the Soviet Union. Abruptly, we find him occupied with the new world of relativity, with a semipopular book and the first volume of a serious text. His most famous contribution was the paper "On the curvature of space" published in the *Zeitschrift für Physik* in 1922, which was at that time one of the main forums for relativity papers. Here he derived an exact solution for $\dot{R}$, the time derivative of the radius of curvature of the Universe. He showed that the Einstein and de Sitter worlds were particular static cases, and if R varied, it must increase. Friedmann demonstrated that R depended on density, so for a critical

density, a closed Universe exists; and he estimated that for a mass of 5×10^{21} Suns and $\Lambda = 0$ (the famous Einstein "cosmological constant") the age of the Universe is 10^{10} years, not a bad estimate for 1922! Einstein soon objected to the concept of a nonstationary world, writing in the journal that the solution did not satisfy the field equations. Friedmann immediately responded with a letter and sent documents to Berlin with Yu. A. Krutkov, who met Einstein in Leiden and convinced him; Einstein withdrew his criticism in *Zeitschrift für Physik* in May 1923. The retraction won Friedmann instant fame in the Soviet Union as "the man who had proved Einstein wrong".

The fame of the Friedmann equation was established by its rediscovery by A. G. E. Lemaître and the discovery of Hubble's law. Friedmann became the director of the Main Geophysical Institute and devoted himself to teaching and reorganizing meteorology in the Soviet Union. He also set the national height record with a balloon ascent to 7,400 metres. In 1923, he followed family tradition by abandoning his first wife for a younger woman. This proved his undoing, for the new couple had a religious marriage ceremony in 1925 in the Crimea; on his way back he caught typhoid fever and his illustrious career ended at the age of 37.

This book is full of interesting information, but it is disjointed. We never get to know Friedmann, who remains an icon. The authors present a lot of detail on relativity, but little on the problems of Communism, although it is pointed out that none of those who followed Friedmann in relativity survived the purges of 1937 (except his student G. Gamow, who fled from the Soviet Union). But the book ends on a triumphant note, with the resurgence of Friedmann's field in the Soviet Union, under the leadership of Ya. B. Zeldovich. □

Harold Zirin is at the Big Bear Solar Observatory, California Institute of Technology, Pasadena, California 91125, USA.

New in physics

Selected Works of Yakov Borisovich Zeldovich. Volume 2: Particles, Nuclei, and the Universe edited by J. P. Ostriker. Princeton University Press, \$69.50 (hbk).

Appropriating the Weather: Vilhelm Bjerknes and the Construction of a Modern Meteorology by R. M. Friedman. Cornell University Press, \$18.95.

The Ghost in the Atom edited by P. C. W. Davies and J. R. Brown. CUP, £5.95, \$9.95. See *Nature* **324**, 420 (1986).

PI in the Sky: Counting, Thinking, and Being by John D. Barrow. Penguin, £7.99. See *Nature* **360**, 376 (1992).

Immunological modulation and evasion by helminth parasites in human populations

Rick M. Maizels, Don A. P. Bundy, Murray E. Selkirk, Deborah F. Smith & Roy M. Anderson

Helminth parasites are highly prevalent in human communities in developing countries. In an endemic area an infected individual may harbour parasitic worms for most of his or her life, and the ability of these infections to survive immunological attack has long been a puzzle. But new techniques are starting to expose the diverse mechanisms by which these agents modulate or evade their hosts' defences, creating a dynamic interaction between the human immune system and the parasite population.

SEVENTY per cent of the global population lives in the developing world, where helminths are one of the commonest types of infection and one third of the population harbours intestinal helminths such as *Ascaris lumbricoides* and *Trichuris trichiura*¹. A child born in an area where intestinal nematodes, schistosome flukes or filarial worms are endemic can expect to harbour worms for most of his or her life owing to repeated exposure to infection and an inability to develop protective immunity.

Infectious agents can be divided into microparasites and macroparasites². The former include viruses, bacteria and protozoa and have direct reproduction within the host, often at high rates. Macroparasites include the helminths and do not multiply within the definitive host, but reproduce sexually to produce transmission stages that egress through faeces or urine, or via a biting arthropod vector. Most microparasitic infections are acute and short-lived and, following recovery, immunity may be protective and even lifelong. In contrast, macroparasites induce long-lasting, chronic infections in which individual organisms persist over long periods.

The ability of helminths to survive in the host for protracted periods despite abundant evidence of immune recognition and attack remains a puzzle of both practical and fundamental significance. Helminth species have complex genomes and elaborate developmental cycles, however, and it would be surprising if they had not evolved mechanisms to combat the host's defence. The growing integration between laboratory and field studies indicates that there is a wide array of mechanisms by which the parasite evades or modulates host immunological attack. A key facet of parasite persistence appears to be enhancement of down-regulatory arms of the immune system to induce tolerance or anergy in the face of repeated infection. This review examines recent epidemiological, immunological, molecular and theoretical studies of the factors that result in persistent infection. Understanding these factors will be central to the development of effective vaccines and better drugs, as well as shedding light on how the existence of these parasites has shaped the evolution of the immune system.

Biological features of helminth infections

In humans, the most striking features of parasitic helminths are long-term persistence within the host³, the ability to elicit protective immunity only after many years or even decades of exposure⁴, complex developmental cycles often involving stage-specific antigens⁵, aggregated distribution in human communities where a minority of people harbour the majority of worms⁶, and the occurrence of individuals predisposed to heavy infection^{7,8}. Helminths cause morbidity rather than mortality, with disease severity typically related to worm burden. The persistent nature of infection, with progressive accumulation of par-

asites, can induce chronic sequelae such as hepatosplenic disease induced by *Schistosoma mansoni*⁹, lymphoedema and elephantiasis induced by lymphatic filariasis and blindness caused by *Onchocerca volvulus*¹. Long-term infection, particularly if associated with poor nutrition, can lead to impaired physical and cognitive development in children with high worm loads¹⁰. Disease may itself result from host immunological responses mediating damage to sites at which parasites have accumulated^{9,11}.

Microparasites often ensure their persistence in human communities by antigenic variation and diversity. Although helminths maintain genetic variability through sexual reproduction (evidenced by the emergence of drug-resistant strains), their generation of new antigen variants is unlikely to be rapid enough because the generation time of the parasite is not much less than that of the host. Genetic heterogeneity within the parasite population may be in part maintained by genetic variation in the human community, where a few hosts predisposed to heavy infection provide a refuge in which parasites can survive and reproduce against a background of deficient immunological defences. Information on genetic heterogeneity in helminth populations, however, remains limited¹².

Transmission between hosts is an uncertain process and, as a result, helminths tend to produce very large numbers of transmission stages (a mature female *Ascaris* can produce 200,000 eggs per day). However, life-cycle strategies vary between species and, as in most organisms, life expectancy of the mature worm, rates of egg or larval production and transmission success (the magnitude of the basic reproductive rate, R_0) are correlated. Long-lived filarial species produce offspring at a lower rate than short-lived intestinal nematodes. Following curative chemotherapy, rates of reinfection are low with filarial worms such as *O. volvulus* (life expectancy 8–10 yr), moderate for schistosome flukes (3–5 yr) and very rapid for intestinal nematodes such as *A. lumbricoides* (1 yr)¹³. Life history strategy and transmission success therefore determine the ideal interval between rounds of mass chemotherapy designed to suppress infection to very low levels¹⁴.

Epidemiological patterns

The epidemiological picture of persistent helminth infection is set by the long parasite lifespan, and by repeated reinfection throughout the host's own lifetime. Estimates of life expectancy of endoparasites residing in the gut, mesenteric veins or the lymphatic system are crude, and depend on studies of the rate of decline of prevalence and intensity of infection in groups of people who cease to be exposed to infection (in the absence of chemotherapy) (Fig. 1a and b). The observed patterns of decay can be translated into a measure of parasite lifespan on the basis of the frequency distribution of parasite numbers per person¹⁵,

which is typically well described by the negative binomial probability distribution, irrespective of the community studied (Fig. 1d). The negative binomial model describes aggregated or clumped distributions (where the variance in worm numbers is greater than the mean) in which most people harbour few worms and a few hosts hold a large number of parasites. For many helminths, a very tight relationship exists between the fraction infected (prevalence) and average worm load (intensity) (Fig. 1c)¹⁶ and measures of the degree of aggregation (measured inversely by the parameter k) are remarkably constant for a given infection over very diverse populations. Aggregation becomes less marked as people age, reflecting either a slow build-up of specific immunity⁴ or variation in susceptibility to infection over time¹³. Taking studies on both infection rate decay and parasite distribution into account, the estimates of life expectancies listed in Table 1 suggest long-term persistence in the host. Surprisingly, persistence in habitats such as the blood or lymphatic system appears longer than in the gut. The long-term survival of tissue-dwelling schistosomes and filarial parasites, despite their most intimate contact with the immune system, suggests that they have evolved particularly effective mechanisms of immune evasion or modulation.

Repeated infection of human populations is revealed by cross-sectional studies that chart changes in worm loads across each age class, and by longitudinal cohort studies in which rates of reinfection following chemotherapy are monitored (Fig. 2a). In the absence of acquired immunity or heterogeneity in exposure, the pattern of change is described by an immigration-death model in which the average intensity rises with age until reaching an equilibrium at the point at which the per capita rate of host infection, Λ , is balanced by the net mortality of the mature

parasite, μ (ref. 3). Rates of reinfection reflect the magnitude of the basic reproductive rate, R_0 , a measure of the average number of offspring produced by one female worm (in the case of dioecious species) that attains reproductive maturity. Reinfection is much faster for *Ascaris* (mean burdens of which can return to precontrol levels within one year of treatment)¹⁸ than for *S. mansoni*^{4,17} or filarial worms¹⁸. All age classes reacquire infection following treatment, but in schistosomiasis^{17,19} older age classes reacquire infection at slower rates than younger groups. This reflects elements both of age-related exposure patterns, and of acquired immunity (integrating the experience of past infection accumulated as the host ages in an environment where repeated exposure takes place). Water contact patterns in schistosome-endemic areas are age-dependent²⁰, as are geophagic habits favouring geohelminth transmission (reflected by the presence of silica in children's faeces)²¹. As a result, the change in mean worm burden with age for many helminths follows a convex curve in which the age of peak intensity depends on various factors, including the life expectancy of the parasite (in younger age classes of children for short-lived intestinal nematodes, and in teenagers or young adults for schistosomes and filarial worms, respectively), the intensity of transmission (earlier in high transmission areas and vice versa), and on behavioural factors (age-related exposure to infection). Interestingly, the pattern of change with age tends to be more peaked or convex in high-transmission areas than in low ones, supporting the hypothesis that acquired immunity develops slowly in a manner related to accumulated past exposure to infection (Fig. 2b)^{22,23}.

Evidence for immune-mediated resistance

Three concepts of importance in immunity to helminth infections

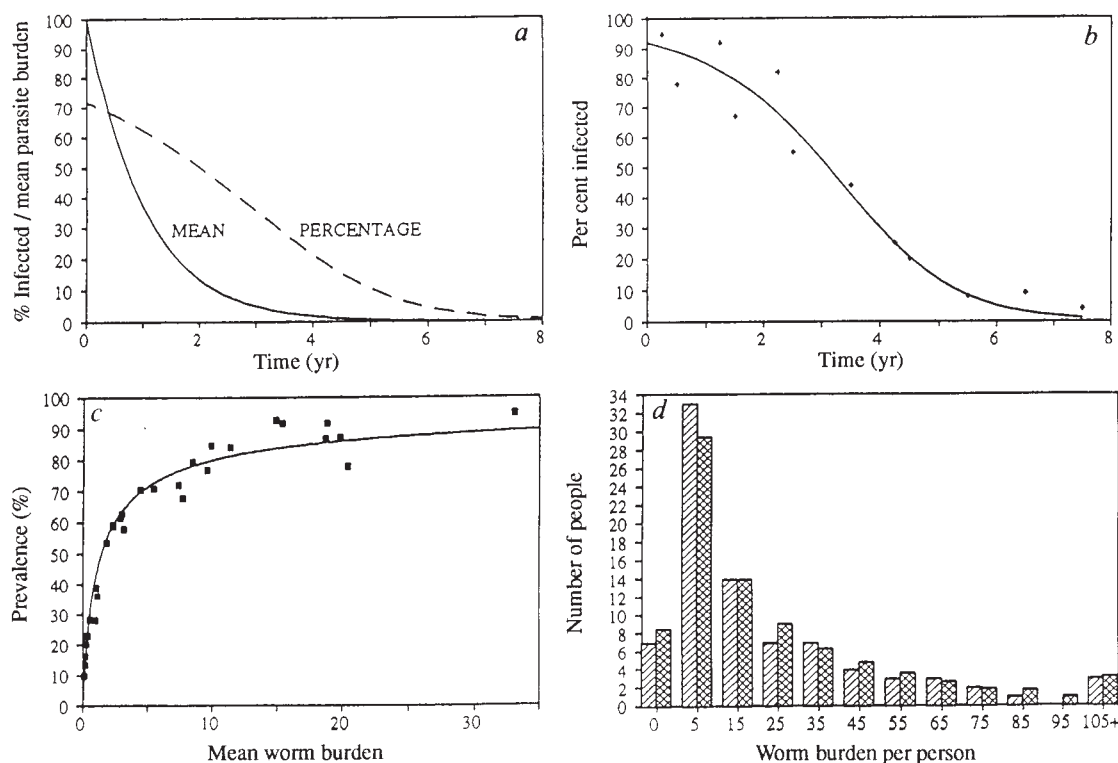


FIG. 1 a, The predicted pattern of decay over time in the prevalence and intensity of infection with a parasite, the distribution of which within the host population is aggregated or overdispersed and whose exposure to infection ceases at time $t=0$. b, The decay in prevalence of *Wuchereria bancrofti* infection within a sample of people no longer exposed to infection. Crosses are observed values¹⁴⁰, and the solid line is the prediction of a simple model of decay¹³. c, The nonlinear relationship between the prevalence of infection (% infected) and the mean intensity of infection (mean worm burden) for *Ascaris lumbricoides*. The solid

line is the relationship predicted by the negative binomial probability model (to describe aggregated distributions) and the squares denote observed values from a wide variety of studies in different Asian and African countries¹⁶. d, The frequency distribution of the number of *Ancylostoma duodenale* per person in a study in India (hatched) compared with the expected values (cross-hatched) of the best-fit negative binomial probability model (mean, 24.54 worms per person; $k=0.618$ (ref. 7)).

TABLE 1 Epidemiological characteristics of some of the principal helminth infections in humans*

Basic parameters of adult helminth parasites in the human host					
Parasite	Common name of parasite (disease)	Average lifespan (yr) [†]	Size of adult (mm)	Location of adult parasite	Taxon
<i>Enterobius vermicularis</i>	Pinworm	<1	10	Colon	Nematode
<i>Trichuris trichiura</i>	Whipworm	1–2	30	Caecum	Nematode
<i>Ascaris lumbricoides</i>	Roundworm	1–2	200	Small intestine	Nematode
<i>Necator americanus</i>	Hookworm	2–3	10	Small intestine	Nematode
<i>Ancylostoma duodenale</i>	Hookworm	2–3	12	Small intestine	Nematode
<i>Schistosoma mansoni</i>	Blood fluke (Bilharzia)	3–5	20 [‡]	Mesenteric veins	Trematode
<i>Schistosoma haematobium</i>	Bladder fluke	3–5	30 [‡]	Vesical plexus	Trematode
<i>Wuchereria bancrofti</i>	Filarial worm	3–5	90 [‡]	Lymphatics	Nematode
<i>Onchocerca volvulus</i>	(River blindness)	8–10	400 [‡]	Subcutaneous	Nematode

Rates of egg production by mature parasites in the human host			Average period from arrival in the human host to maturation		
Parasite	Rate of egg production (per female per day)	Role of egg in transmission of disease	Parasite	Mode of entry	Maturation delay (d)
<i>Ascaris lumbricoides</i>	>200,000	Embryonates in soil to become infective egg if ingested by mouth	<i>Schistosoma haematobium</i>	Percutaneous	21–28
<i>Ancylostoma duodenale</i>	10,000–20,000	Hatches in soil to produce skin-penetrating infective larva (L3)	<i>Schistosoma mansoni</i>	Percutaneous	25–30
<i>Schistosoma haematobium</i>	500–3,000	Hatches in water to release miracidium; this invades snail	<i>Ancylostoma duodenale</i>	Percutaneous	28–50
<i>Schistosoma mansoni</i>	100–300	intermediate host and reproduces to yield >1,000 aquatic cercariae able to penetrate human skin	<i>Ascaris lumbricoides</i>	Oral	50–80
			<i>Trichuris trichiura</i>	Oral	50–84
			<i>Onchocerca volvulus</i>	Blackfly bite	365 +

Degree of parasite aggregation in human communities §	
Parasite	k value
<i>Enterobius vermicularis</i>	0.3–0.4
<i>Trichuris trichiura</i>	0.2–0.4
<i>Ascaris lumbricoides</i>	0.2–0.9
<i>Necator americanus</i>	0.03–0.4
<i>Schistosoma mansoni</i>	0.03–0.5
<i>Wuchereria bancrofti</i>	0.6–0.7

* See ref. 13 for sources of estimates.

[†] Lifespan estimates are rough approximations owing to practical difficulties inherent in estimation. For further details on helminth biology, see refs 131, 143.

[‡] Female.

§ Measured inversely by parameter *k* of the negative binomial probability distribution of worm numbers per host (for *k* values greater than 5, the distribution tends to the Poisson (random) form).

are age-related acquisition of resistance, concomitant immunity, and predisposition to infection.

■ There is strong evidence for age-dependent acquisition of immunity in humans to the major helminth disease of schistosomiasis^{4,17,19,23}; a similar pattern is seen in human filariasis^{24,25}. Exposure is a major determinant of age-related changes in schistosomiasis (which is dependent on water contact), but not in filariasis (which is mosquito-borne). At all ages, specific antibody and cellular responses targeted at parasite antigens abound, with some measures correlating positively with worm burden^{26–28} (Fig. 2c) and others with duration of exposure.

■ In concomitant immunity, hosts are protected against newly invading larvae while tolerating an established adult worm load. Originally described in experimental infections with schistosomes²⁹, cestodes³⁰ and nematodes³¹ in rodent hosts, the limited evidence for concomitant immunity in humans must be dissociated from effects due to the natural rate of parasite population turnover and reinfection within a single host^{17,19,20,25}, or from pathophysiological sequelae of infection which may change the niche available to later waves of parasites³². Concomitant immunity would imply stage-specific expression of antigens by larval helminths, and/or an ability of adult worms to subvert or counteract those immune effector mechanisms that kill immature forms. In either case, the infective forms appear valid targets for vaccine development on the basis that human immune systems appear able to reduce their rate of establishment, given some considerable experience of past exposure.

■ Predisposition to heavy or light infection is an important epidemiological factor, now established for *Ascaris*, *Trichuris*,

hookworms and schistosomes^{7,33,34} by studies of reinfection following chemotherapy. Genetic, behavioural, environmental or nutritional factors may singly or in combination cause this effect. Predisposition is more easily demonstrated in children than adults, and is evident, typically, in the 20–40% of individuals who show a significant positive correlation between pre- and post-treatment worm burdens⁸.

The highly aggregated distribution of parasite numbers per person (Table 1) may reflect differences in exposure, in susceptibility to infection or in the ability to mount effective immunological responses. Recent theoretical work on the dynamics of the Th1–Th2 interaction in helminth infections and on the effects of parasite mediated modulation of effector mechanisms suggest that early exposure in infancy (or even in the womb via maternal experience of infection³⁵) may determine predisposition to heavy or light infection where high exposure generates anergy with elevated Th2 responses concomitant with high worm loads (Fig. 2d)³⁶. Theory suggests that the quantitative details of the host's ability to recognize and respond to particular parasite antigens is of great importance in determining whether the host is persistently susceptible or develops protective immunity. Incremental changes in the ability to recognize or respond to a particular antigen can alter the state to which the dynamical system is attracted.

Different responses to individual antigens among subjects with similar exposure experience (Fig. 3) is likely to be influenced by genetic³⁷, nutritional³⁸ and behavioural³⁹ components. Recent data demonstrate individual variability, in antibody recognition of particular parasite antigens⁴⁰, and in T-cell responses depend-

ing upon the host major histocompatibility complex (MHC) genotype⁴¹. Similar work in animal models suggests that both MHC and non-MHC genes can be critical to the course of infection^{37,42}. In human populations, the combination of these variable factors is undoubtedly a major determinant of parasite aggregation and predisposition.

To tease out the factors that control observed patterns of infection within communities, epidemiological study must turn more and more to detailed longitudinal studies of parasitological and immunological changes in individuals within the community. Recent field studies of this type are beginning to identify patterns which provide important clues to what, amongst many measurable responses to parasite antigens, may determine the build up of acquired immunity which acts to reduce the intensity of infection in older individuals. For example, in studies of the gut-dwelling worm *T. trichiura*, comparison of age-dependent isotype responses and the age-profile of infection intensity in two endemic communities with very different levels of transmission, suggest that parasite-specific secretory IgA responses best reflect the accumulated past experience of infection, whereas serum IgG

isotypes simply correlate positively with current worm burden^{43,44}. For the blood-dwelling schistosome flukes in humans (*S. mansoni*²⁸ and *S. haematobium*²⁶), circulating specific and nonspecific serum IgE correlates both positively with accumulated past experience of infection, and negatively with reinfection rates after drug treatment.

Immune responses and effector mechanisms

How does specific immunological resistance to helminth parasites operate? Some information can be drawn from epidemiological correlations, but our understanding rests heavily upon animal model systems and *in vitro* studies (Table 2). In dissecting the elaborate web of immune responses mounted to helminth parasites, one must note that, as argued by Mitchell³⁰, individual mechanisms rarely operate in isolation in the natural host, and the most efficient expression of host resistance is likely to be effected by a multicomponent response. Helminths have evolved an array of modulations and interference strategies to match the defence systems of their preferred host, and there is considerable functional overlap between independent immune cell types.

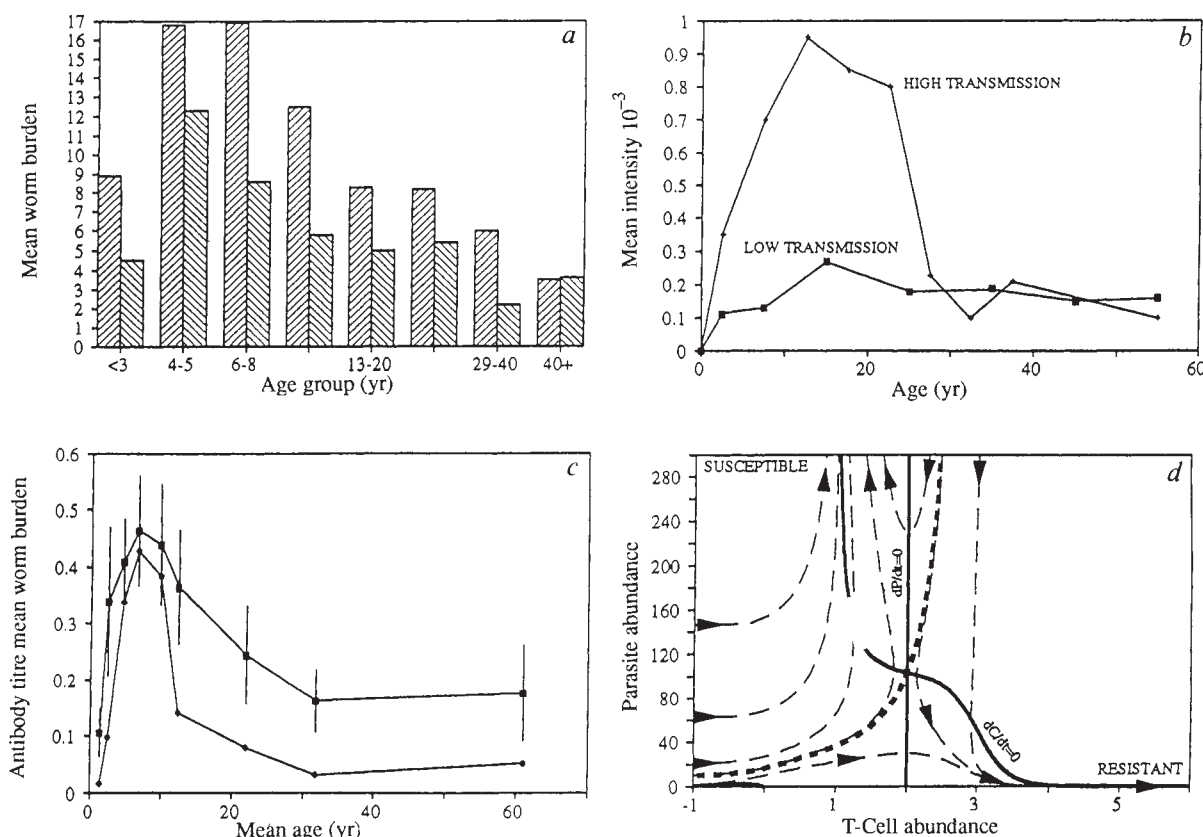


FIG. 2 a, A horizontal age-stratified survey of the intensity of infection (worm burden) with *Ascaris lumbricoides* in a rural community in India before chemotherapeutic treatment (January 1984; hatched bars) and post-treatment (November 1984; cross-hatched bars) showing rapid rates of reinfection where worm burdens across all age classes approach pre-control levels after 11 months of exposure¹⁴¹. b, Horizontal age-stratified surveys of the intensity of infection (eggs per g of faeces) with *Schistosoma mansoni* in areas of low and high transmission. Note the convex curve in the high-transmission area and the monotonic approach to a plateau with increased age in the low-transmission area. c, Horizontal age-stratified study of the intensity of infection (worm burden) with *Trichuris trichiura* in a community in St Lucia, West Indies (diamonds) and age-related changes in *T. trichiura* antigen-specific IgG antibody titres (95% confidence limits around squares) in sera taken from the study population. Note that antibody titres rise and fall across age classes in a manner related to worm burden¹⁴². It is unclear

whether humoral responses are protective. d, Mathematical models of the dynamic interaction between Th1 lymphocytes and helminth parasite antigens reveal complex patterns of dynamical behaviour where early experience of infection may determine whether the host acquires resistance or remains susceptible as a result of immunological unresponsiveness generated by suppressed Th1 cell abundance. The graph shows a phase plane of the relationships between helminth parasite abundance (P) and Th1 cell abundance (C). The solid lines represent isoclines ($dP/dt=0$ and $dC/dt=0$). The heavy broken line represents a stable manifold (boundary) between two areas of attraction (that is, susceptibility or immunity) and the lightly broken lines represent numerical solutions of the model with varying initial conditions (that is, values of P and C at time $t=0$). Note that the high exposure to the parasite in early life (when C is small) can result in susceptibility (unresponsiveness), whereas low exposure may result in the slow acquisition of resistance³⁶.

These factors militate against deletion analyses in which single components are removed from an experimental system, particularly when a non-natural host is involved.

The prime candidates for effective immune mechanisms against helminth organisms are the humoral and cellular components most evident during infection. The most distinctive of these features is a dramatic elevation of serum IgE antibody levels in humans⁴⁵ and rodents⁴⁶: the nematode *Nippostrongylus brasiliensis* raises total serum IgE levels in rodents 100-fold in 14 days, and 10-fold higher again on a secondary infection. This reflects a polyclonal activation, as responses to other antigens introduced before the helminth infection are also the subject of a strong IgE response⁴⁷. The potentiation of IgE (and in humans of IgG4) is mediated by a soluble cytokine, interleukin (IL)-4 (refs 48, 49), the main source of which is the Th2 helper cell subset.

Eosinophilia is equally marked in helminth infections: it has long been known that in human helminthiasis more than 50% of circulating white blood cells may be eosinophil granulocytes⁵⁰, in contrast to their normal level of 2–5%. Similarly, in rodent infections, circulating eosinophil numbers increase 10 to 30-fold⁵¹, a phenomenon blocked by antibodies to another Th2 lymphokine, IL-5 (ref. 52). Helminthiasis also provoke proliferation of mast cells (IgE-receptor bearing cells which release pharmacologically active mediators), particularly in the mucosa of parasitized intestinal tissue³¹. This mastocytosis is promoted by IL-3 and other cytokines⁵³. Lymphocytes from helminth-infected, eosinophilic patients show enhanced IL-4 and IL-5 production, again suggesting that these cytokines are directly amplified as a result of infection⁵⁴.

Whether the abundance of IgE and eosinophils actually protect the host against helminth parasites is a matter of some contention. In human schistosomiasis, both elevated eosinophilia^{55,56} and specific IgE responses^{26,28} correlate quantitatively with resistance to reinfection after chemotherapy, whereas in filariasis, patent microfilariemia is associated with a depressed

IgE response⁵⁷. However, IgE-mediated hypersensitivity (due to mast cell discharge of active mediators) is linked with pathology rather than resistance in human gut trichuriasis infections⁵⁸. Helminth parasites can be killed *in vitro* by IgE-mediated mechanisms, involving platelets, macrophages and eosinophils (Table 2)⁴⁵. More evidence for a protective role of IgE is offered by studies on rats infected with filarial worms⁵⁹ and *T. spiralis*⁶⁰. Although protection of rats against *S. mansoni* implicates IgE⁴⁵, mice depleted of IgE and eosinophils can be satisfactorily vaccinated⁶¹. In conclusion, IgE probably contributes to protection against tissue helminths in humans and rats, if not always in mice.

Eosinophils certainly kill larval schistosomula *in vitro*, primarily by deposition of toxic granule proteins⁶² in synergy with reactive oxygen intermediates (ROI; for example, free superoxide and hydroxyl radicals, and hydrogen peroxide)⁶³ (Table 2). *In vivo*, there is conflicting evidence on the ability of eosinophil-depleted animals to curtail a *S. mansoni* infection^{61,64}, but hypoeosinophilic animals have compromised immunity to migrating larvae of *Strongyloides venezuelensis*⁶⁵ and *Trichinella spiralis*⁵¹. In both instances immunity to gut-living stages was unabated, demonstrating that protective mechanisms in this locale are quite distinct.

Efficient expulsion of intestinal helminths generally involves a parasite-specific induction phase which recruits a more catholic effector phase with many common elements³¹. During expulsion, a highly oedematous local inflammatory response occurs, characterized by mastocytosis, epithelial shedding, increased gut permeability and mucus hypersecretion⁶⁶. These and other cell types may contribute inflammatory lipid mediators^{66,67} and damaging oxygen free radicals⁶⁸. Inhibition of mastocytosis by the nematode *Heligmosomoides polygyrus* permits both this species, and other concurrent parasites, to remain in the gut⁶⁹. Permeability changes introduce plasma effector molecules, and some of these combine with mucus in order to immobilize parasites before expulsion (Table 2)^{31,70}. An important point is that the same effector mechanisms that attack intestinal worms are probably responsible for the concurrent gut pathology⁶⁸. This may well be true for many tissue-dwelling parasites, and represents a major barrier to vaccine development.

In many settings, the role of lymphokine-activated macrophages appears well established. The lymphokines that activate murine macrophages are interferon- γ (IFN- γ), acting in synergy with tumour necrosis factor (TNF)⁷⁴, both products of the Th1 subset of helper lymphocytes. Schistosomula are killed by murine^{71,72} and human⁷³ lymphokine-activated macrophages (Table 2) but not by reactive oxygen intermediates produced by phagocytes stimulated through Fc receptors⁷⁴. Activated murine macrophages kill schistosomula by the production of nitric oxide⁷² so that, in the mouse at least, resistance is strongly associated with IFN- γ production, which is in turn primarily generated by Th1 subset cells⁷⁵.

Th1 and Th2 helper T cells

The two subsets of differentiated helper T cells, Th1 and Th2, are characterized by the contrasting profiles of cytokines they secrete. The Th1 products (including IFN- γ and TNF- β) are inflammatory mediators and selectively activate macrophages, whereas the Th2 cytokines (such as IL-4, IL-5 and IL-10) stimulate B-cell and eosinophil development and antibody production. Moreover, these subsets are reciprocally cross-inhibitory⁷⁵, so that one cell type will gain at the expense of the other; this may explain the profoundly unbalanced Th1/Th2 ratio in most helminth infections⁷⁶. The elevated eosinophilia, IgE and IgG4 expression in chronic human helminthiasis suggest that Th2 cells predominate, and indeed higher IL-4 and lower IFN- γ production is reported in peripheral T cells from heavily infected individuals^{54,77}, and human T-cell clones reactive to antigens from the nematode *Toxocara canis* show a pure Th2 profile⁷⁸. However, the helper subset status in infection may not be irre-

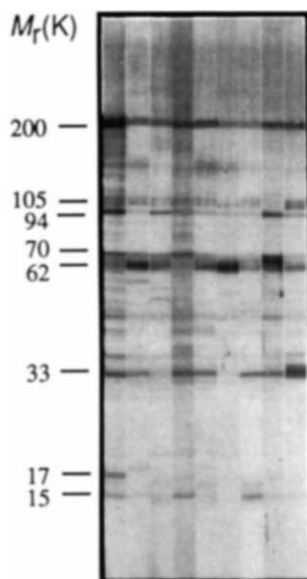


FIG. 3 Evidence that different people (of similar ages and with similar backgrounds of exposure to infection) vary in their ability to recognize and respond to particular antigens. Sera from 9 individual adult patients with high levels (>100 mf ml^{-1}) of *Wuchereria bancrofti* microfilariae in nocturnal blood samples were used to probe a western blot of infective larval antigens taken from a closely related filarial parasite, *Brugia malayi*²⁴. No correlation was observed between the expression of a particular antibody specificity and parasitological factors such as microfilarial density, resistance to infective larvae, and age. The sizes of major components, such as myosin (200K), are indicated.

TABLE 2 Effector mechanisms known to act against helminths

Effector cells	Effector molecules	Likely mechanism	Cytokine (Th subset) dependence	Parasite	Host	Ref.
IFN- γ -activated macrophages	—	Production of nitric oxide	IFN- γ (Th1)	<i>S. mansoni</i>	Mouse	72
Monocytes	IgE	Reactive oxygen intermediates	IL-4 (Th2)	<i>S. mansoni</i>	Human	45,144
Eosinophils	IgE	ECP, MBP, ROI	IL-4 + IL-5 (Th2)	<i>S. mansoni</i>	Human	45,145
Eosinophils	IgG1 or IgG3	ECP, MBP, ROI	—	<i>S. mansoni</i>	Human	132
Neutrophils	IgG or IgM	ROI and granule proteins	—	<i>Dirofilaria immitis</i>	Dog	133
Platelets	IgE	ROI and/or prostanoids	IL-4 (Th2)	<i>S. mansoni</i>	Human, rat	45,134
Platelets	CRP	ROI and/or prostanoids	—	<i>S. mansoni</i>	Rat	135
Mucosal mast cells	IgE	Histamine and/or prostanoids	IL-3 + IL-4 (Th2)	<i>T. spiralis</i>	Rat	31,67
—	IgG + complement	Membrane attack complex	—	<i>T. taeniaeformis</i>	Mouse	136
—	Mucus + antibodies	Trapping and exclusion	—	<i>N. brasiliensis</i>	Rat	31,70

Only selected examples of each mode of helminth killing are presented for brevity. For a review of cited references, see ref. 61. Details of *S. mansoni* are given in Table 1. *Dirofilaria immitis* is the mosquito-transmitted filarial nematode of dogs, heartworm. *Nippostrongylus brasiliensis* is a strongylid nematode in the gut of rodents which follows a hookworm-like life cycle. *Taenia taeniaeformis* is a cestode found in the adult stage in the cat intestine, forming cysts in intermediate rodent hosts on ingestion of eggs. Abbreviations: ECP, eosinophil cationic protein; MBP, major basic protein; ROI, reactive oxygen intermediates (hydrogen peroxide, hydroxyl radicals, hypohalous acids, singlet oxygen and superoxide); prostanoids include hydroxyeicosatetraenoic acids, leukotrienes, prostacyclin, prostaglandins and thromboxane.

vocable; in murine schistosomiasis, an early Th1-line response is superseded by a stronger Th2 response after egg production⁷⁹ or following repeated immunization with attenuated larvae⁸⁰.

At issue is whether Th1 cells are generally protective against helminth parasites. Although the singly immunized mouse model for schistosomiasis suggests that Th1 is the protective cell type for larvicidal responses⁷⁹, evidence from infections with the whipworm, *Trichuris muris*, argues cogently against drawing a general conclusion of this nature. In this system, susceptible strains of inbred mice show Th1 activation, measured by high IFN- γ production, whereas resistant mice produce little IFN- γ and elevated IL-5, symptomatic of a strong Th2 response⁸¹. Moreover, anti-larval (concomitant) immunity to the mouse nematode *Heligmosomoides polygyrus* is ablated by anti-IL-4 treatment⁸², and can be passively transferred by IgG1 antibodies⁸³, which are potentiated by IL-4 (ref. 48). However, in *T. spiralis*, expression of Th1-mediated properties, such as high IFN- γ and serum IgG2a, have been associated with fast expulsion of the gastrointestinal adult worms⁸⁴, confirming that the critical T-cell population may differ in each parasite combination and in each tissue site. Thus, there is as yet no evidence for a common anti-helminth immune mechanism; as parasites have evolved multifarious ways of evading appropriate steps in various immune cascades, generalizations must await data from a sufficiently broad range of host-parasite combinations. In this context, patterns of T-cell cytokine production in human helminth infection^{54,77,85} will illuminate both the role of T-cell subsets in parasitic disease, and more generally establish the extent to which the human T-cell response conforms to the highly polarized Th1/2 distinction found in mice^{86,87}.

Parasite defence strategies

Infectious organisms have evolved a diverse array of specific and general adaptations to evade the host immune system^{88,89}. The defence strategies used by helminths^{89,90}, constrained by replication rate and extracellular locale, can be divided into three classes: avoiding initial induction of damaging immune responses; compromising selected arms of the immune system; and disabling the short-range offensives mounted by various effector mechanisms. (A summary of helminth molecules involved in these defences is given in Table 3.)

■ Some devices exploited by helminths to block induction have been well described: molecular mimicry and uptake of host antigens^{29,91} are two means by which a helminth may confer 'self' status upon itself, whereas secretion of the hapten phosphorylcholine may be a diversionary tactic⁵. Schistosomes absorb host MHC⁹¹, conceivably as defence against natural killer (NK) cells (which react to aberrant expression of MHC), but also take up decay-accelerating factor (conferring protection from complement-mediated attack⁹²), contrapsin (an antithrombotic serum serine protease inhibitor⁹³), and low-density lipoproteins (possibly causing the progressive loss of antibody binding sites with maturation⁹⁴). In these instances, uptake of host molecules is quintessentially parasitic, to subvert host functions to the well-being of the parasite.

Because these strategies offer only limited scope for expression of products necessary for the parasitic mode of life, more directed means of regulating immunity are required. The extracellular niche of helminths makes the immune system dependent upon the exogenous pathway for antigen presentation. However, macrophages from schistosome⁹⁵ and *H. polygyrus*⁹⁶-infected mice are incompetent antigen presenters. Possibly the quantity of polysaccharides and glycoconjugates produced may interfere with antigen processing⁹⁷; secretions containing phosphorylcholine from nematodes may act to inhibit activation⁹⁸. Antigen processing generally requires cysteinyl and aspartyl proteases such as cathepsins B and D⁹⁹. Filarial worms release a cystatin-like molecule¹⁰⁰ which may block cathepsin B, whereas antigen processing in B cells can be blocked by an aspartic protease inhibitor from *Ascaris lumbricoides*¹⁰¹. Other protease inhibitors act in the extracellular milieu: taeniaestatin, a 19.5K serine protease inhibitor from metacestode larvae of *Taenia taeniaeformis*, inhibits chemotaxis by neutrophils, as well as T-cell proliferation and production of IL-2 (ref. 102). Interference with signalling and antigen presentation may not only block a response, but also induce a state of immunological unresponsiveness, or tolerance¹⁰³, in which there can be long-term assimilation of parasites into their host environment. The dynamics of a chronic infection and persistent antigen would favour maintenance of the tolerant state^{36,104}.

■ An equally effective strategy may be to neutralize only those components of the immune system that pose a real danger. Para-

TABLE 3 Helminth products that may interact with host immune systems

Parasite	Product	Comments	Ref.
Immunoglobulin-cleaving proteases			
<i>Dirofilaria immitis</i>		On parasite surface	107
<i>Fasciola hepatica</i>		On parasite surface	106
<i>Schistosoma mansoni</i>		Surface	105
Protease inhibitors			
<i>Echinococcus granulosus</i>	12K Ag B	Secreted into cyst fluid	109
<i>Onchocerca volvulus</i>	Cystatin	On parasite surface	100
<i>Taenia-taeniaformis</i>	Taeniaestatin	IL-2/chemotaxis inhibitor	102
Antioxidants			
<i>Brugia pahangi</i>	GPX	Surface/secreted	122
<i>Onchocerca volvulus</i>	SOD	Secreted	137
<i>Schistosoma japonicum</i>	GST	On parasite surface	123
<i>Schistosoma mansoni</i>	GST	On parasite surface	125
Other factors			
<i>Brugia malayi</i>	Prostaglandin E2	Mf secreted	111
<i>Heligmosomoides polygyrus</i>	Immunosuppressant	Adult worm associated	96
<i>Nippostrongylus brasiliensis</i>	PAF hydrolase	Secreted	110
<i>Schistosoma mansoni</i>	Immunosuppressant	Secreted	138
<i>Taenia solium</i>	Paramyosin	C1-binding; secreted	139
<i>Taenia taeniaeformis</i>	Sulphated proteoglycan	C3-consuming; secreted	108
<i>Taenia taeniaeformis</i>	Prostaglandin E2	Larval secretions	112

Abbreviations: SOD, superoxide dismutase; PAF, platelet-activating factor; GPX, glutathione peroxidase; GST, glutathione-S-transferase; C1 and C3, first and third components of complement; IL-2, interleukin-2. Details on *O. volvulus* and *S. mansoni* are given in Table 1, and on *D. immitis*, *N. brasiliensis* and *T. taeniaeformis* in Table 2. *Brugia malayi* and *B. pahangi* are filarial nematodes closely related to *W. bancrofti* (Table 1). *Echinococcus granulosus* is a cestode of dogs, transmitted through eggs found in soil, which causes hydatid cysts in humans. *Fasciola hepatica* is the liver fluke, parasitizing sheep and cattle and transmitted through an intermediate snail host. *Heligmosomoides polygyrus* is a gastrointestinal nematode of mice, with a direct faecal-oral cycle. *Schistosoma japonicum* is a blood fluke with a similar life cycle to that of *S. mansoni*. *Taenia solium* is the tapeworm acquired by humans from cysts in infected pork; the same species causes cysticercosis if eggs are ingested.

sites can directly or indirectly block the effects of antibody with surface or secreted proteases capable of degrading host immunoglobulin molecules¹⁰⁵⁻¹⁰⁷, while complement function is compromised in tapeworm cysts by highly sulphated proteoglycans which activate complement in the fluid phase¹⁰⁸. Host inflammatory reactions are also interrupted by helminths. The cestode *Echinococcus granulosus* secretes an elastase inhibitor which blocks neutrophil attraction by C5a or platelet-activating factor (PAF)¹⁰⁹. PAF may be neutralized directly through cleavage by secreted acetylhydrolases, as documented for *N. brasiliensis*¹¹⁰. Filarial¹¹¹ and taeniae¹¹² parasites use endogenous and exogenous arachidonic acid to produce and release prostanoids (prostaglandin and PGE₂), which in addition to these anti-inflammatory properties may also inhibit T-cell proliferation.

Selective inhibition of T-cell subsets is evidenced in filariasis in which proliferative T cells show antigen-specific anergy, but antibody responses remain intact^{85,113}; this has been interpreted as Th1-subset anergy¹¹⁴. High levels of IgE and IgG4 are produced during helminth infections. Co-expression of these two isotypes may be beneficial as IgG4 can block IgE-mediated allergic responses¹¹⁵. Protective mechanisms may also be modulated by the IgG4:IgE ratio, which tends to be high in heavily infected individuals in schistosomiasis²⁶ and filariasis⁵⁷. Antibodies of the IgM and IgG2 isotypes (directed against carbohydrate determinants) also block responses to schistosomula by granulocytes which are guided by antibodies of other isotypes in the antibody-dependent cell-mediated cytotoxicity reaction¹¹⁶. The presence of these blocking antibodies is correlated with susceptibility to reinfection following chemotherapy of a human population, whereas expression of IgG1 and IgG3 antibodies to the same epitopes correlate with resistance.

■ The third layer of the helminth strategy involves a set of short-term defences to ward off the many forms of intense attack which may be mounted by the host immune system. Physical effects include the thickening of the schistosome tegument during maturation¹¹⁷, and the corresponding inaccessibility of the nem-

atode hypodermal plasma membrane beneath a thick extracellular cuticle. The increasing invulnerability of maturing helminths may contribute towards the operation of concomitant immunity. Many nematodes have a loose surface coat which can be readily sloughed off under immune attack¹¹⁸, and *Fasciola* juvenile flukes show a similar rapid turnover of surface glycocalyx¹¹⁹.

To counteract the oxidative burst of activated host leukocytes, helminths express surface or secreted antioxidant enzymes¹²⁰ such as superoxide dismutase¹²¹, glutathione peroxidase (GPX)¹²² and glutathione S-transferase (GST)^{123,124}. GST is localized in the external tegument of schistosomes¹²⁵, and successfully immunizes rats against challenge¹²⁴. The major role of these enzymes may be to protect surface membranes against peroxidation; it is notable that GST/GPX expression, and parasite resistance to oxidants, each progressively rise through the schistosome life cycle¹²⁶.

Discussion

Successful helminth parasites combine specific molecular strategies to combat the threat of immediate immune attack, with quantitative and dynamic properties which permit long-term establishment within an individual host (by tolerance induction, for example) and within the host population (such as by overdispersion in 'wormy' individuals). Recent molecular and immunological advances have developed new insights and opportunities for the analysis of macroparasite transmission and immunity in human communities. There is a growing emphasis in epidemiological studies on the dynamic interaction between parasite populations and the immune system, on delineating temporal changes in parasite levels and immunological responses, and on recording variability between human hosts. In human filariasis, for example, the rate of increase in infection can now be measured with assays for parasite antigen circulating in the bloodstream²⁵, and the host response can be monitored more precisely by measuring antibodies of a single isotype¹²⁷ or to a single parasite antigen¹²⁸.

The structural and developmental sophistication of macropar-

asites is clearly reflected in the multiplicity of pathways by which parasites interact with the host immune system. It is often argued that the Th2 subset (and, in particular, IgE) has evolved to help combat helminth infection. If so, it may also be that extant helminth parasites have evolved mechanisms that allow them to counteract Th2 responses. The success of these parasites (as indicated by their high prevalences in the developing world) suggests a very dynamic evolution where continued selection enables the parasite to evade or modulate the immunological attack in at least some fraction of the host population (the 'wormy' individuals).

Theoretical studies of parasite mediated modulation of immune responses and the dynamic interaction between Th1 and Th2 responses to helminth infections have generated new hypotheses to explain heterogeneity and predisposition based on the induction of specific immunological non-responsiveness (tolerance or anergy) to parasite antigens through high exposure very early in life³⁶. These ideas require testing in areas of endemic infection by 'immuno-epidemiological studies of infants and children (including those in whom neonatal tolerance may have resulted from maternal infection³⁵), and of laboratory systems. Key to these hypotheses are the influence of the parasite in disrupting the regulation of helper (responsive) and suppressor (unresponsive) activities in a multicomponent immune response.

Heterogeneity in parasite burdens, where a small fraction of the community harbours the majority of the parasites, also has implications for parasite evolution. If those predisposed to heavy infection are less immunologically competent to resist repeated infection (due to genetic or nutritional factors), they therefore provide a refugium for the evolution of new variants via sexual reproduction. Transmission of these variants to the majority of immunocompetent hosts will select for viable organisms with enhanced ability to modulate or evade immunological responses.

The many opportunities presented by advances in molecular biology for the study of immunity to helminths are paralleled by the dangers of an ever-expanding descriptive literature¹²⁹. The population ecology of the immune system is undoubtedly a very nonlinear world: future research will need to emphasize the measurement and quantification of the rate parameters that control humoral and cellular responses to parasite antigens, and the functional dependencies between parasite and immune system variables that determine the dynamics of the interaction¹³⁰. The Th1/Th2 balance or imbalance is a good example of this need. Simple theory points to two attractors in this dynamic interaction of enhanced or suppressed responsiveness (Fig. 2d), separated by an unstable boundary where small quantitative changes in particular parameters of the immune system can switch the host from anergy and unresponsiveness to immunity. The need for the new molecular tools is beginning to be appreciated by the field epidemiologist, but perhaps immunologists have not always appreciated the need for field and mathematical approaches.

To understand patterns of infection, reinfection and response, molecular techniques must be integrated into long-term studies of exposure and infection. Future field research will then be able to monitor the population dynamics of immunological responses, as well as to conduct longitudinal studies in individual patients that chart changes in particular modes of the immune response in relation to infection and the slow acquisition of immunity in the majority of those exposed¹¹⁶. To help interpret field observations, laboratory studies are required with animal models that centre on the dynamics of repeated exposure as the host ages, and which examine long-term presentation of antigens, as opposed to the short exposure to abrupt pulses of parasites or defined antigens upon which current models of immune responsiveness have been built. The marked heterogeneity in worm loads within human communities, and the evidence for predisposition to heavy or light infection is a help not a hindrance in such research. It is becoming increasingly apparent that heterogeneity not only reflects differences in exposure, but

also host genetic factors involved in epitope selection and immune effector function. Work on laboratory rodent/helminth systems suggests that there may be host genes that determine predisposition³⁷, and the association between HLA genotype and infection intensity is now being actively analysed.

The all-too-apparent complexity of the interaction between helminth parasites and the human immune system (our appreciation of which has been magnified by new molecular and immunological tools) may deter rather than encourage the choice of these systems as models for the study of infection and immunity. This is a crucial challenge, however, because understanding how these organisms persist in the host, despite a diversity of specific and nonspecific immunological defences directed at them, is likely to provide essential new insights, not only into the functioning of the human immune system, but also into the design of vaccines and therapies to combat antigenically diverse infectious agents. □

The authors are at the Wellcome Research Centre for Parasitic Infections, Departments of Biology (R.M.M., D.A.P.B. and R.M.A.) and Biochemistry (M.E.S. and D.F.S.), Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BB, UK.

- Warren, K. S. & Mahmoud, A. A. F. *Tropical and Geographic Medicine* (McGraw-Hill, New York, 1984).
- Anderson, R. M. & May, R. M. *Nature* **280**, 361–367 (1979).
- Anderson, R. M. & May, R. M. *Adv. Parasit.* **24**, 1–101 (1985).
- Butterworth, A. E., Fulford, A. J. C., Dunne, D. W., Ouma, J. H. & Sturrock, R. F. *Phil. Trans. R. Soc. B* **321**, 495–511 (1988).
- Maizels, R. M. & Selkirk, M. E. in *The Biology of Parasitism* (eds Englund, P. T. & Sher, A.) 285–308 (Liss, New York, 1988).
- Anderson, R. M. *Trans. R. Soc. Trop. Med. Hyg.* **80**, 686–696 (1986).
- Schad, G. A. & Anderson, R. M. *Science* **228**, 1537–1540 (1985).
- Keymer, A. & Pagel, M. in *Hookworm Disease: Current Status and New Directions* (eds Schad, G. A. & Warren, K. S.) 177–209 (Taylor and Francis, London, 1990).
- Warren, K. S. in *Bailliere's Clinical Tropical Medicine and Communicable Diseases* (ed. Mahmoud, A. A. F.) 301–313 (Bailliere, London, 1987).
- Nokes, C., Grantham-McGregor, S. M., Sawyer, A. W., Cooper, E. S. & Bundy, D. A. P. *Proc. R. Soc. B* **247**, 77–81 (1992).
- Ottesen, E. A. *Rev. Infect. Dis.* **7**, 796–801 (1985).
- Nadler, S. A. *Parasit. Today* **3**, 154–155 (1987).
- Anderson, R. M. & May, R. M. *Infectious Diseases of Humans. Dynamics and Control* 1–757 (Oxford Univ. Press, Oxford, 1991).
- Anderson, R. M. & Medley, G. F. *Parasitology* **90**, 629–660 (1985).
- Anderson, R. M. & May, R. M. *Nature* **297**, 557–563 (1982).
- Guyatt, H. L., Bundy, D. A. P., Medley, G. F. & Grenfell, B. T. *Parasitology* **101**, 139–143 (1990).
- Butterworth, A. E. *et al. Trans. R. Soc. Trop. Med. Hyg.* **79**, 393–408 (1985).
- Bundy, D. A. P., Grenfell, B. T. & Rajagopalan, P. K. in *Immunoparasit. Today* (eds Ash, C. & Gallagher, R. B.) A71–A75 (Elsevier, Cambridge, 1991).
- Wilkins, H. A., Blumenthal, U. J., Hagan, P., Hayes, R. J. & Tulloch, S. *Trans. R. Soc. Trop. Med. Hyg.* **81**, 29–35 (1987).
- Chandiwna, K. K., Woolhouse, M. E. J. & Bradley, M. *Parasitology* **102**, 73–83 (1991).
- Wong, M. S. *Trans. R. Soc. Trop. Med. Hyg.* **85**, 89–91 (1990).
- Anderson, R. M. & May, R. M. *Nature* **315**, 493–496 (1985).
- Woolhouse, M. E. J., Taylor, P., Matanhire, D. & Chandiwna, S. K. *Nature* **351**, 757–759 (1991).
- Day, K. P., Gregory, W. F. & Maizels, R. M. *Parasite Immun.* **13**, 277–290 (1991).
- Day, K. P., Grenfell, B., Spark, R., Kazura, J. W. & Alpers, M. P. *Am. J. Trop. Med. Hyg.* **44**, 518–527 (1991).
- Hagan, P., Blumenthal, U. J., Dunn, D., Simpson, A. J. G. & Wilkins, H. A. *Nature* **349**, 243–245 (1991).
- Bundy, D. A. P., Lillywhite, J. E., Didier, J. M., Simmons, I. & Bianco, A. E. *Parasite Immun.* **13**, 629–638 (1991).
- Dunne, D. W. *et al. Eur. J. Immun.* **22**, 1483–1494 (1992).
- Smithers, S. R. & Terry, R. J. *Ann. N.Y. Acad. Sci.* **160**, 826 (1969).
- Mitchell, G. F. *Immunology* **38**, 209–223 (1979).
- Miller, H. R. P. *Vet. Immun. Immunopath.* **6**, 167–259 (1984).
- Wilson, R. A. *Parasit. Today* **6**, 354–358 (1990).
- Haswell-Elkins, M. R., Elkins, D. B. & Anderson, R. M. *Parasitology* **95**, 323–328 (1987).
- Tingley, G. A. *et al. Trans. R. Soc. Trop. Med. Hyg.* **82**, 448–452 (1988).
- Lammie, P. J., Hitch, W. L., Walker Allen, E. M., Hightower, W. & Eberhard, M. L. *Lancet* **337**, 1005–1006 (1991).
- Schweitzer, N. & Anderson, R. *Proc. R. Soc. B* **247**, 107–112 (1992).
- Wakelin, D. in *Genetics of Resistance to Bacterial and Parasitic Infection* (eds Wakelin, D. M. & Blackwell, J. M.) 153–232 (Taylor and Francis, London, 1988).
- Bundy, D. A. P. *Parasitology* **95**, 623–635 (1987).
- Bundy, D. A. P. & Medley, G. F. *Parasitology* **104**, 105–119 (1992).
- Kennedy, M. W. *Parasit. Today* **5**, 316–324 (1989).
- Hirayama, K. *et al. Nature* **327**, 426–430 (1987).
- Wassom, D. L., Wakelin, D., Brooks, B. O., Krco, C. J. & David, C. S. *Immunology* **51**, 625–631 (1984).
- Needham, C. S. *et al. Parasitology* **105**, 275–283 (1992).
- Needham, C. S. & Lillywhite, J. E. *Trans. R. Soc. Trop. Med. Hyg.* (in the press).
- Capron, A. & Dessaint, J. P. A. *Rev. Immun.* **3**, 455–476 (1985).
- Ogilvie, B. M., Smithers, S. R. & Terry, R. J. *Nature* **209**, 1221–1223 (1966).
- Jarrett, E. E. & Miller, H. R. P. *Prog. Allergy* **31**, 178–233 (1982).
- Finkelman, F. D. *et al. Rev. Immun.* **8**, 303–333 (1990).
- Lundgren, M. *et al. Eur. J. Immun.* **19**, 1311–1315 (1989).
- Brown, T. R. *J. exp. Med.* **3**, 315–347 (1898).

51. Grove, D. I., Mahmoud, A. A. F. & Warren, K. S. *J. exp. Med.* **145**, 755–759 (1977).
52. Coffman, R. L., Seymour, B. W. P., Hudak, S., Jackson, J. & Rennick, D. *Science* **245**, 308–310 (1989).
53. Madden, K. B. *et al. J. Immun.* **147**, 1387–1391 (1991).
54. Mahanty, S., Abrams, J. S., King, C. L., Limaye, A. P. & Nutman, T. B. *J. Immun.* **148**, 3567–3571 (1992).
55. Sturrock, R. F. *et al. Trans. R. Soc. Trop. Med. Hyg.* **77**, 363–371 (1983).
56. Hagan, P., Wilkins, H. A., Blumenthal, U. J., Hayes, R. J. & Greenwood, B. M. *Parasite Immun.* **7**, 625–632 (1985).
57. Kurniawan, A. *et al. J. Immun.* **150**, 3941–3950 (1993).
58. Cooper, E. S. *et al. Lancet* **338**, 1104–1107 (1991).
59. Gusmão, R. de, Stanley, A. M. & Ottesen, E. A. *Expt Parasit.* **52**, 147–159 (1981).
60. Dessein, A. J., Parker, W. L., James, S. L. & David, J. R. *J. exp. Med.* **153**, 423–436 (1981).
61. Sher, A., Coffman, R. L., Hiery, S. & Cheever, A. W. *J. Immun.* **145**, 3911–3916 (1990).
62. Butterworth, A. E. *Adv. Parasit.* **23**, 143–235 (1984).
63. Yazdanbakhsh, M., Tai, P.-C., Spry, C. J. F., Gleich, G. J. & Roos, D. *J. Immun.* **138**, 3443–3447 (1987).
64. Mahmoud, A. A. F., Warren, K. S. & Peters, P. A. *J. exp. Med.* **142**, 805–813 (1975).
65. Korenaga, M. *et al. Immunology* **72**, 502–507 (1991).
66. Moqbel, R. & MacDonald, A. J. in *Parasites: Immunity and Pathology. The Consequences of Parasitic Infection in Mammals* (eds Behnke, J. M.) 249–282 (Taylor and Francis, London, 1990).
67. Moqbel, R. *et al. Immunology* **60**, 425–430 (1987).
68. Smith, N. C. *Parasit. Res.* **75**, 423–438 (1989).
69. Dehlawi, M. S., Wakelin, D. & Behnke, J. M. *Parasite Immun.* **9**, 187–194 (1987).
70. Miller, H. R. P., Huntley, J. F. & Wallace, G. R. *Immunology* **44**, 419–429 (1981).
71. Esparza, I., Männel Ruppel, A., Falk, W. & Krammer, P. H. *J. exp. Med.* **166**, 589–594 (1987).
72. James, S. L. & Glaven, J. J. *Immun.* **143**, 4208–4212 (1989).
73. Cottrell, B. J., Pye, C., Glauert, A. M. & Butterworth, A. E. *J. Cell Sci.* **94**, 733–741 (1989).
74. Scott, P., James, S. & Sher, A. *Eur. J. Immun.* **15**, 553–558 (1985).
75. Mosmann, T. R. & Coffman, R. L. *A. Rev. Immun.* **9**, 145–173 (1989).
76. Sher, A. F. & Coffman, R. L. *A. Rev. Immun.* **10**, 385–409 (1992).
77. Zwingenberger, K., Hohmann, A., Cardoso de Brito, M. & Ritter, M. *Scand. J. Immun.* **34**, 243–251 (1991).
78. Del Prete, G. *et al. J. clin. Invest.* **88**, 346–350 (1991).
79. Pearce, E. J., Caspar, P., Grzych, J.-M., Lewis, F. A. & Sher, A. *J. exp. Med.* **173**, 159–166 (1991).
80. Caulada-Benedetti, Z., Al-Zamel, F., Sher, A. & James, S. J. *Immun.* **146**, 1655–1660 (1991).
81. Else, K. J. & Grecnis, R. K. *Immunology* **72**, 508–513 (1991).
82. Urban, J. F. Jr, Katona, I. M., Paul, W. E. & Finkelman, F. D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5513–5517 (1991).
83. Pritchard, D. I., Williams, D. J. L., Behnke, J. M. & Lee, T. D. *G. Immunology* **49**, 353–365 (1983).
84. Pond, L., Wassom, D. L. & Hayes, C. E. *J. Immun.* **143**, 4232–4237 (1989).
85. Yazdanbakhsh, M. *et al. J. infect. Dis.* **167**, 925–931 (1993).
86. Romagnani, S. *Immun. Today* **12**, 256–257 (1991).
87. Wierenga, E. A. *et al. J. Immun.* **147**, 2942–2949 (1991).
88. Gooding, L. R. *Cell* **71**, 5–7 (1992).
89. Dessaint, J.-P. & Capron, A. R. in *Immunology and Molecular Biology of Parasitic Infections* 3rd edn (ed. Warren, K. S.) 87–99 (Blackwell Scientific, Boston, 1993).
90. Behnke, J. M., Barnard, C. J. & Wakelin, D. *Int. J. Parasit.* **22**, 861–907 (1992).
91. Sher, A. & Moser, G. *Am. J. Path.* **102**, 121–126 (1981).
92. Pearce, E. J., Hall, B. F. & Sher, A. *J. Immun.* **144**, 2751–2756 (1990).
93. Modha, J., Parikh, V., Gauldie, J. & Doenhoff, M. *J. Parasitology* **96**, 99–109 (1988).
94. Chiang, C.-P. & Caulfield, J. P. *Am. J. Path.* **134**, 1007–1018 (1989).
95. Stadecker, M. J., Kamisato, J. K. & Chikunguwo, S. M. *J. Immun.* **145**, 2697–2700 (1990).
96. Pritchard, D. I., Ali, N. M. H. & Behnke, J. M. *Immunology* **51**, 633–642 (1984).
97. Leyva-Cobian, F. & Unanue, E. R. *J. Immun.* **141**, 1445–1450 (1988).
98. Lal, R. B., Kumaraswami, V., Steel, C. & Nutman, T. B. *Am. J. Trop. Med. Hyg.* **42**, 56–64 (1990).
99. Diment, S. J. *Immun.* **145**, 417–422 (1990).
100. Lustigman, S., Brotman, B., Huima, T. & Prince, A. M. *Molec. Biochem. Parasit.* **45**, 65–76 (1991).
101. Bennett, K. *et al. Eur. J. Immun.* **22**, 1519–1524 (1992).
102. Leid, R. W., Suquet, C. M., Bouwer, H. G. A. & Hinrichs, D. J. *J. Immun.* **137**, 2700–2702 (1986).
103. Schwartz, R. H. *Science* **248**, 1349–1356 (1990).
104. Ramsdell, F. & Fowlkes, B. J. *Science* **257**, 1130–1134 (1992).
105. Auriault, C., Ouassii, M. A., Torpier, G., Eisen, H. & Capron, A. *Parasite Immun.* **3**, 33–41 (1981).
106. Chapman, C. B. & Mitchell, G. F. *Vet. Parasit.* **11**, 165–178 (1982).
107. Tamashiro, W. K., Rao, M. & Scott, A. L. *J. Parasit.* **73**, 149–154 (1987).
108. Hammerberg, B. & Williams, J. F. *J. Immun.* **120**, 1033–1037 (1978).
109. Shepherd, J. C., Aitken, A. & McManus, D. P. *Molec. Biochem. Parasit.* **44**, 81–90 (1991).
110. Blackburn, C. C. & Selkirk, M. E. *Immunology* **75**, 41–46 (1992).
111. Liu, L. X., Serhan, C. N. & Weller, P. F. *J. exp. Med.* **172**, 993–996 (1990).
112. Leid, R. W. & McConnell, L. A. *Prostagland. Leukotri. Med.* **11**, 317–323 (1983).
113. Nutman, T. B., Kumaraswami, V. & Ottesen, E. A. *J. clin. Invest.* **79**, 1516–1523 (1987).
114. Maizels, R. M. & Lawrence, R. A. *Parasit. Today* **7**, 271–276 (1991).
115. Ottesen, E. A., Kumaraswami, V., Paranjape, R., Pindexter, R. W. & Tripathy, S. P. *J. Immun.* **127**, 2014–2020 (1981).
116. Butterworth, A. E. *et al. Parasitology* **94**, 281–300 (1987).
117. Dessein, A. *et al. Parasitology* **82**, 357–374 (1981).
118. Blaxter, M. L., Page, A. P., Rudin, W. & Maizels, R. M. *Parasit. Today* **8**, 243–247 (1992).
119. Hanna, R. E. *B. Expl Parasit.* **50**, 103–114 (1980).
120. Callahan, H. L., Crouch, R. K. & James, E. R. *Parasit. Today* **4**, 218–225 (1988).
121. Simurda, M. C., van Keulen, H., Rekosh, D. & LoVerde, P. T. *Expl Parasit.* **67**, 73–84 (1988).
122. Cookson, E., Blaxter, M. L. & Selkirk, M. E. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5837–5841 (1992).
123. Smith, D. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 8703–8707 (1986).
124. Ballou, J. M. *et al. Nature* **326**, 149–153 (1987).
125. Taylor, J. B. *et al. EMBO J.* **7**, 456–472 (1988).
126. Nare, B., Smith, J. M. & Prichard, R. K. *Expl Parasit.* **70**, 389–397 (1990).
127. Kwan-Lim, G.-E., Forsyth, K. P. & Maizels, R. M. *J. Immun.* **145**, 4298–4305 (1990).
128. Dissanayake, S., Xu, M. & Piessens, W. F. *Molec. Biochem. Parasit.* **56**, 269–278 (1992).
129. Maddox, J. *Nature* **363**, 13 (1993).
130. Anderson, R. M. *J. Anim. Ecol.* **60**, 1–50 (1991).
131. Muller, R. *Worms and Disease, a Manual of Medical Helminthology* (Heinemann Medical, London, 1975).
132. Muller, R. in *Immunology and Molecular Biology of Parasitic Infections* 3rd edn (ed. Warren, K. S.) 567–598 (Blackwell Scientific, Boston, 1993).
133. Joseph, M., Capron, A., Butterworth, A. E., Sturrock, R. F. & Houba, V. *Clin. Exp. Immunol.* **33**, 48–56 (1978).
134. Capron, M. *et al. J. Immun.* **132**, 462–468 (1984).
135. Khalife, J. *et al. J. Immun.* **142**, 4422–4427 (1989).
136. Rzepczyk, C. M. & Bishop, C. J. *Parasite Immun.* **6**, 443–457 (1984).
137. Joseph, M., Auriault, C., Capron, A., Vorng, H. & Viens, P. *Nature* **303**, 810–812 (1983).
138. Bout, D. *et al. Science* **231**, 153–156 (1986).
139. Mitchell, G. F. *Adv. Immun.* **28**, 451–511 (1979).
140. Henkle, K. J., Liebau, E., Müller, S., Bergmann, B. & Walter, R. D. *Infect. Immun.* **59**, 2063–2069 (1991).
141. Mazingue, C. *et al. Int. Arch. Allerg. appl. Immun.* **83**, 12–18 (1987).
142. Laclette, J. P. *et al. J. Immun.* **148**, 124–128 (1992).
143. Webber, R. H. S. E. *Asian J. trop. Med. Pub. Hlth* **27**, 59–85 (1975).
144. Elkins, D. B., Haswell-Elkins, M. & Anderson, R. M. *Trans. R. Soc. Trop. Med. Hyg.* **80**, 774–792 (1986).
145. Bundy, D. A. P. *Phil. Trans. R. Soc. Lond. B* **321**, 405–420 (1988).

ACKNOWLEDGEMENTS. We thank the Wellcome Trust for support through the Wellcome Research Centre for Parasitic Infections at Imperial College.

The isotopic compositions and stellar sources of meteoritic graphite grains

Sachiko Amari^{*†}, Peter Hoppe^{*‡}, Ernst Zinner^{*} & Roy S. Lewis[†]

^{*} McDonnell Center for the Space Sciences and the Physics Department, Washington University, St Louis, Missouri 63130, USA

[†] Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA

The isotopic compositions of interstellar graphite grains in meteorites provide a record of the stellar environments in which they formed. The grains have $^{12}\text{C}/^{13}\text{C}$ ratios that range far from the terrestrial value, as well as large variations in $^{14}\text{N}/^{15}\text{N}$ and $^{16}\text{O}/^{18}\text{O}$, and pronounced ^{26}Mg excesses from the decay of extinct ^{26}Al . This isotopic variability indicates several types of stellar source, including asymptotic giant branch stars, Wolf-Rayet stars and novae.

PRIMITIVE meteorites contain interstellar dust grains that formed in circumstellar atmospheres and survived the formation of the Solar System without melting or vaporization^{1,2}. The chemical and isotopic study of these grains thus provides important information on stellar nucleosynthesis and complements direct astronomical observations³. Interstellar grains in meteorites include diamond⁴, silicon carbide^{5,6}, graphite⁷, titanium carbide⁸ and corundum^{9,10}. Although diamonds (mean size 1.6 nm) are too small to be analysed as single grains, the other types of interstellar grains can have diameters in excess of 1 μm and are amenable to single-grain analysis of their chemical and isotopic properties by ion microprobe mass spectrometry^{7,9,14}. Silicon carbide is rich in trace elements, which makes it possible to measure isotopic ratios of C, Si, N, Mg, Ca, Ti, Ba, Nd and Sm in the ion microprobe^{11,15}, and Sr, Ba, Nd and Sm by thermal ionization mass spectrometry^{16–19}, in addition to the noble gases^{11,20,21}.

In contrast, much less information has been obtained to date on interstellar graphite, largely because of its lower abundance in meteorites, the smaller content of trace elements and the more complicated isolation procedure^{7,22}. Here we report C-, N- and O-isotopic data, and inferred $^{26}\text{Al}/^{27}\text{Al}$ ratios on individual graphite grains (0.7–11 μm in size), separated from the Murchison CM2 chondrite. Although both graphite and silicon carbide are expected to condense in similar cooling circumstellar envelopes^{23–25}, the isotopic compositions of these two types of interstellar grains are quite distinct, indicating that they have different origins.

Samples and procedures

Previous studies have shown that the carrier of the exotic noble gas component Ne-E(L)—almost pure ^{22}Ne (ref. 26)—is graphite in the density range 1.75–2.2 g cm^{-3} and of nominal grain size >1 μm (ref. 7). A sample of 114 g of fusion-crust-free chips was treated with HF–HCl to dissolve silicates, then with KOH to remove reactive kerogen and sulphur, and the resulting residue slowly oxidized with $\text{Cr}_2\text{O}_7^{2-}$. After a colloidal extraction to remove diamonds, the residue was separated according to density in sodium polytungstate solution with a continuous density gradient. This procedure revealed distinct bands and four fractions were extracted; KE1 (1.6–2.05 g cm^{-3} , 1.2 p.p.m. of total meteorite). There was some substructure within this density range but no further separation was undertaken, KFA1 (2.05–2.10 g cm^{-3} , 0.22 p.p.m.), KFB1 (2.10–2.15 g cm^{-3} , 0.15 p.p.m.) and KFC1 (2.15–2.20 g cm^{-3} , 0.44 p.p.m.). A detailed description of the separation procedure is given in ref. 22.

All four density fractions contain Ne-E(L) (ref. 27). Differences in the thermal release profiles²⁷ and in the Kr isotopic

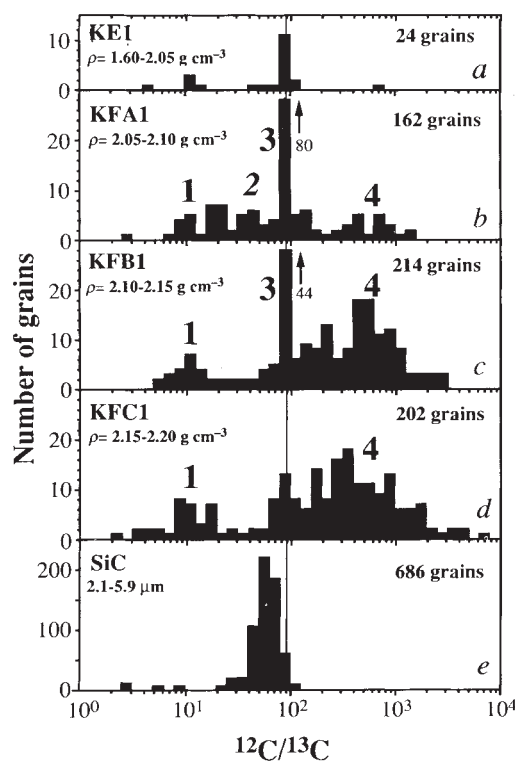


FIG. 1 a–d, Histograms of $^{12}\text{C}/^{13}\text{C}$ ratios in round graphite grains. The abundance of round grains depends on density. It is <10% in fraction KE1 (a), >50% in KFA1 (b) and KFB1 (c) and ~70% in KFC1 (d). The distributions of $^{12}\text{C}/^{13}\text{C}$ indicate the presence of at least three distinct populations: 1 ($3 < ^{12}\text{C}/^{13}\text{C} < 20$), 3 ($87 < ^{12}\text{C}/^{13}\text{C} < 96$) and 4 ($100 < ^{12}\text{C}/^{13}\text{C} < 7,000$). The terrestrial value of 89 is represented by a thin vertical line. In contrast to graphite, $^{12}\text{C}/^{13}\text{C}$ ratios in interstellar silicon carbide grains of comparable size (e) display a much narrower range³¹. Fraction KFA1 seems to have a higher abundance of grains with $^{12}\text{C}/^{13}\text{C}$ ratios between 20 and 60 than the other graphite fractions; this is a region populated by many silicon carbide grains. Furthermore, these grains seem to have distinct morphology. But because of limited statistics we only tentatively postulate a separate population for these graphite grains (population No. 2). The vertical arrows in b and c indicate the numbers of grains with $^{12}\text{C}/^{13}\text{C}$ ratios between 79 and 100; these numbers exceed the vertical ranges of the two plots.

[‡] Present address: Physikalisches Institut, Universität Bern, Switzerland.

compositions²⁸ indicate the presence of at least four separate components. The grains in the graphite samples have different morphologies. Most are round and dense, but others have irregular shapes. The round grains in turn come in two basic types; the first type has smooth surfaces and sometimes display shell-like platy structures, whereas the second type seems to consist of a dense aggregate of small scales. These surface morphologies are also reflected in the internal structure of the grains as revealed by transmission electron microscopy of microtomed grains⁸. The smooth or platy grains show layers of well crystallized graphite (see Fig. 2a of ref. 8), usually enclosing a core of amorphous carbon (T. J. Bernatowicz, personal communication); the scaly spherules appear to consist of concentric layers of smaller units of low graphitization with turbostratic textures (Fig. 2b of ref. 8). Grains of both types sometimes contain small crystals of titanium and other refractory carbides (ref. 8 and T. J. Bernatowicz, personal communication).

As previous work had shown that only the round grains have large anomalies in their C-isotopic compositions²⁹, we concentrated our analysis mostly on those grains (see, however, Fig. 2a). Ion microprobe isotopic analyses of C, N and O were made with a Cs⁺ beam of 2–3 µm diameter by techniques described in refs 11 and 30. Magnesium and Al were measured as positive secondary ions produced by O[−] bombardment. Because the low yield of positive C ions makes locating graphite grains by ion imaging extremely difficult, only a limited number of especially large grains were analysed for their Mg isotopes.

Isotopic compositions

The ¹²C/¹³C ratio in all grains ranges from 2.1 to >7,200. Histograms of this ratio for round grains are plotted in Fig. 1a–d, and may be compared to the distribution of ¹²C/¹³C ratios in interstellar silicon carbide grains of comparable sizes³¹ shown in Fig. 1e. The distributions of ¹²C/¹³C are not uniform but indicate the presence of several populations, which are denoted by arabic numerals in Fig. 1a–d. The isotopic composition of the grains depends on morphology as well as density. This relationship is complex and a detailed description will be reported elsewhere (P.H., S.A., E.Z. and R.S.L., manuscript in preparation). A group with ¹²C/¹³C ratios in the range 3–20 is present in all density fractions (population 1 in Fig. 1). Separates KFA1 and KFB1 show large proportions of grains with ¹²C/¹³C ratios close to the terrestrial value of 89 (population 3). Finally, all separates, except possibly KE1, have isotopically light grains with a wide distribution of ¹²C/¹³C (from 100 to >7,000) (population 4). In separates KFB1 and KFC1 the ¹²C/¹³C ratio does not seem to depend strongly on grain size, but in KFA1 larger (>4 µm) grains tend to be isotopically normal. The isotopically normal grains in KFA1 form a tight cluster (Fig. 2b) with ¹²C/¹³C ratios between 87 and 96. The average ratio of this cluster is 91.4 ($\delta^{13}\text{C}_{\text{PDB}} = [({}^{13}\text{C}/{}^{12}\text{C})_{\text{sample}}/({}^{13}\text{C}/{}^{12}\text{C})_{\text{PDB}} - 1] = -26\%$). It would be tempting to ascribe a Solar System origin to grains with normal C were it not for the fact that two of them, one from KFA1 and one from KFB1, have inferred ²⁶Al/²⁷Al ratios of 0.061 and 0.086, which are much higher than the value of 5×10^{-5} for the early Solar System. Also, seven (four from KFA1, three from KFB1) have anomalous N and three (one from KFA1, two from KFB1) have anomalous Si.

In view of the extremely large range in ¹²C/¹³C ratios and the fact that nucleosynthetic processes affecting the C ratios are expected also to affect the N ratios, the rather limited range in ¹⁴N/¹⁵N is puzzling (Fig. 2). Grains from KE1 show the largest variations (Fig. 2a). Data for KFC1 grains are not plotted because, due to their small size (63% are smaller than 2 µm) and low N contents, errors on the ¹⁴N/¹⁵N ratio are very large and, within these errors, almost all grains have normal N. There is a tendency for grains from all separates to have ¹⁵N excesses rather than deficits relative to the Solar System (see Fig. 2). A possible explanation for the relatively normal N in the grains is exchange or dilution of the original N with isotopically normal N, either in

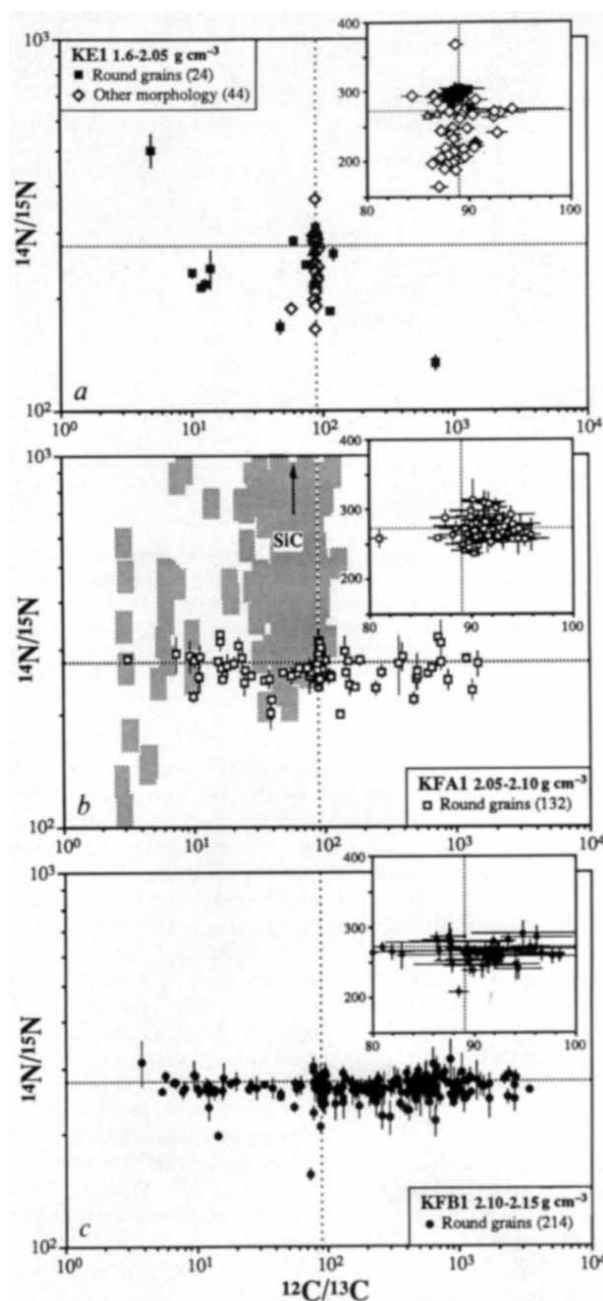


FIG. 2 Isotope ratios in graphite grains; fraction KE1 (a), fraction KFA1 (b) and fraction KFB1 (c). Whereas the ¹²C/¹³C ratios of graphite grains vary over a large range, the range of ¹⁴N/¹⁵N ratios is more limited. KFC1 grains are not plotted because of large errors in the ¹⁴N/¹⁵N ratios. b shows also the C- and N-isotopic compositions of interstellar SiC grains³¹ (shaded areas). Thin dotted lines indicate terrestrial ratios for C (89) and N (272). The numbers of grains plotted for the different density fractions are not the same as those in the histograms of Fig. 1 because sometimes a smaller number of grains was analysed for their C and N isotopic compositions than for their C isotopes only. The insets show the C- and N-isotopic ratios of grains with ¹²C/¹³C ratios close to the terrestrial value of 89. Note the linear scales of the inset plots.

the Solar System or during chemical processing in the laboratory (possibly treatment with NH₃?). If this were the case, original data points in Fig. 2 would have been much further away from the line of solar N and would move toward this line due to the exchange/dilution process.

In Fig. 2a, we also plot the compositions of KE1 grains with non-round morphologies. Their C-isotopic ratios are close to normal (inset), but many have anomalous N, mostly ¹⁵N

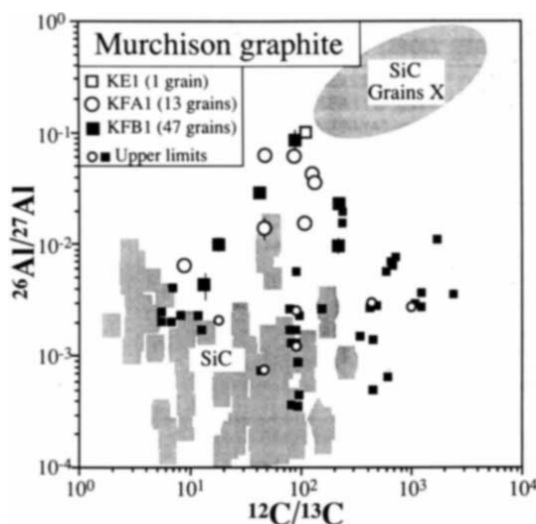


FIG. 3 The $^{26}\text{Al}/^{27}\text{Al}$ ratios of graphite grains, inferred from ^{26}Mg excesses, peak for $^{12}\text{C}/^{13}\text{C}$ ratios in the vicinity of the solar ratio (89). Also shown (shaded areas) are the compositions of interstellar silicon carbide grains³¹. Grains X constitute a rare (~1%) subclass of silicon carbide grains characterized by excesses of ^{12}C , ^{15}N , ^{28}Si , and ^{49}Ti and by extremely large inferred $^{26}\text{Al}/^{27}\text{Al}$ ratios^{31,33}.

excesses, on average larger than those found in round KFA1 and KFB1 grains. Because N isotopic anomalies can be generated by chemical reactions³², it is not possible to prove a circumstellar origin for these grains, and we shall not discuss them any further. In Fig. 2b, we indicate for comparison (shaded area) the corresponding region for most interstellar SiC grains³¹. As with their C-isotopic compositions (Fig. 1), there is also a clear distinction between graphite and silicon carbide grains with regard to their N-isotopic ratios (Fig. 2). Even if the measured N-isotopic ratios are not the original ones, exchange/dilution by normal N cannot make data points cross the line of the solar $^{14}\text{N}/^{15}\text{N}$ ratio and most graphite grains with isotopically heavy C plot below this line, in contrast to most silicon carbide grains.

The difference between graphite and silicon carbide grains extends to the $^{26}\text{Al}/^{27}\text{Al}$ ratios inferred from excesses in ^{26}Mg (Fig. 3). The $^{26}\text{Al}/^{27}\text{Al}$ ratios in graphite range up to 0.1, larger than in silicon carbide grains except for a very rare and special class of grains termed grains X (refs 31, 33). Whereas (except for a few grains with $^{12}\text{C}/^{13}\text{C}=55$) most silicon carbide grains have the highest $^{26}\text{Al}/^{27}\text{Al}$ ratios at very low $^{12}\text{C}/^{13}\text{C}$ ratios, the highest $^{26}\text{Al}/^{27}\text{Al}$ ratios of graphite grains are shown by grains of roughly normal C-isotopic ratios (Fig. 3).

One of the most interesting isotopic signatures of graphite grains are those of oxygen (Fig. 4). Most grains (total number analysed is: KE1, 6; KFA1, 17; KFB1, 14; KFC1, 51) have normal O, but seven show large ^{18}O excesses (up to a factor of 13). Note that grains with ^{18}O excesses have also ^{15}N excesses, and that two (the only ones large enough for Mg isotopic analysis) have large $^{26}\text{Al}/^{27}\text{Al}$ ratios. There is even a rough correlation between the ^{18}O and ^{15}N excesses for six grains, and only the KFC1 grain lies outside this trend. A correlation between the ^{18}O excess and the $^{12}\text{C}/^{13}\text{C}$ ratio is less clear, but among the grains with O anomalies the grain with the largest ^{18}O excess has the highest $^{12}\text{C}/^{13}\text{C}$ ratio (722). The six grains from KE1 and KFA1 with ^{18}O excesses have normal $^{17}\text{O}/^{16}\text{O}$ ratios³⁴.

Stellar sources

The isotopic compositions of the interstellar graphite grains of this study show a high degree of complexity and cannot be explained by a single type of stellar source. The condition of $\text{C}/\text{O} > 1$, necessary for the condensation of graphite, is satisfied in several stellar sources. (1) Carbon stars; these are asymptotic giant branch (AGB) stars during their thermally pulsing phase.

(2) Wolf-Rayet stars; these are massive mass-losing stars. During earlier stages in their evolution their surfaces are dominated by N resulting from H-burning (WN phase); later on the products of He-burning, either C (WC phase) or O (WO phase) are exposed on their surfaces. (3) Novae (explosive H-burning on the surface of a white dwarf). (4) The He-burning shell of a pre-supernova.

Although carbon stars are believed to be the most prolific suppliers of carbonaceous dust to the interstellar medium³⁵, the C-isotopic compositions of most graphite grains do not agree with either those observed in C-stars^{36–38} (where most $^{12}\text{C}/^{13}\text{C}$ values are between 30 and 80), or with those of interstellar silicon carbide (Fig. 1). The $^{12}\text{C}/^{13}\text{C}$ ratios in population 1 are much smaller. Also, N tends to be isotopically normal or slightly heavy relative to Solar System material (Fig. 2), in contrast to silicon carbide. Grains with $20 < ^{12}\text{C}/^{13}\text{C} < 80$, whose C-isotopic ratios are in the range shown both by C-stars and by the bulk of silicon carbide, have mostly heavy N (Fig. 2b). A strong argument for the presence of at least a certain number of grains from AGB stars comes from noble gas measurements. All density fractions of graphite contain Kr with s-process signature²⁸ and most of the s-process elements are believed to be produced in the He-shell of thermally pulsing AGB stars³⁹ with $^{13}\text{C}(\alpha, n)^{16}\text{O}$ as the neutron source. However, s-process nucleosynthesis (capture of neutrons at a slow rate relative to β -decay rates) can also take place during He-burning in Wolf-Rayet stars^{40,41}, and the Kr-isotopic ratios of graphite²⁸ indicate that both AGB and Wolf-Rayet stars are likely to have made contributions (S. Amari, R. S. Lewis & E. Anders, manuscript in preparation). Unfortunately, we do not know the fraction of grains from such sources, their C- and N-isotopic compositions or why they appear to be different from those of interstellar silicon carbide.

Wolf-Rayet stars during the WN phase are expected to have small $^{12}\text{C}/^{13}\text{C}$ ratios, close to the equilibrium value of the CNO cycle (~3.5), on the surface^{42,43} and in some models⁴³ $\text{C} > \text{O}$. The problem with such stellar sources is that the $^{14}\text{N}/^{15}\text{N}$ ratio is expected to be of the order of 10^4 , definitely higher than solar.

Population 3 grains with isotopically normal C are enigmatic. The $^{12}\text{C}/^{13}\text{C}$ ratios of KFA1 grains (in the range 80–100) cluster around a value characteristic of organic terrestrial material, raising the suspicion that we are dealing with contamination. However, the anomalous N, Mg and Si isotopic compositions in a certain fraction of these grains (5 out of 80 with $85 < ^{12}\text{C}/^{13}\text{C} < 96$ in KFA1, and 6 out of 35 in KFB1) indicate that at least some, if not all, are of interstellar origin. (These fractions are lower limits because not all grains with isotopically normal C had their N, Mg and Si isotopics measured.)

The question of origin seems clearer for population 4. Wolf-Rayet stars during the WC phase, when the products of He-burning appear at the surface, are carbon-rich and have high $^{12}\text{C}/^{13}\text{C}$ ratios^{42,43}. An alternative is the He-burning shell of massive stars that precede a type II supernova^{44,45}. In this layer some He has been consumed, making it ^{12}C -rich, whereas ^{13}C and ^{14}N have been destroyed. What makes both models even more attractive are the ^{18}O excesses measured in graphite grains. During H-burning ^{14}N is the most abundant nuclear species of the CNO elements. At the onset of He-burning, ^{14}N is converted to ^{18}O by $^{14}\text{N}(\alpha, \gamma)^{18}\text{F}(\beta, \nu)^{18}\text{O}$ before being further changed into ^{22}Ne by $^{18}\text{O}(\alpha, \gamma)^{22}\text{Ne}$. In his model for massive stars (25 and 40 $M_{\odot}$ solar masses) Maeder⁴³ obtains a large ^{18}O abundance at the WN–WC transition. Similarly, in the He-burning shell of pre-supernova stars, the O-isotopic composition is ^{18}O -rich and ^{17}O -poor^{44,45}. Excesses of ^{15}N in grains with excess of ^{18}O might be due to the more rapid destruction of ^{14}N ; at temperatures typical of He-burning, the destruction of ^{15}N (by $^{15}\text{N}(\alpha, \gamma)^{19}\text{F}$) proceeds slower than the destruction of ^{14}N (ref. 46). Additional evidence comes from abundances of ^{26}Al . This radionuclide is produced by H-burning, but is still present at the surface of Wolf-Rayet stars during the WN–WC transition and during the WC phase. Prantzos *et al.*⁴² predict a $^{26}\text{Al}/^{27}\text{Al}$ ratio

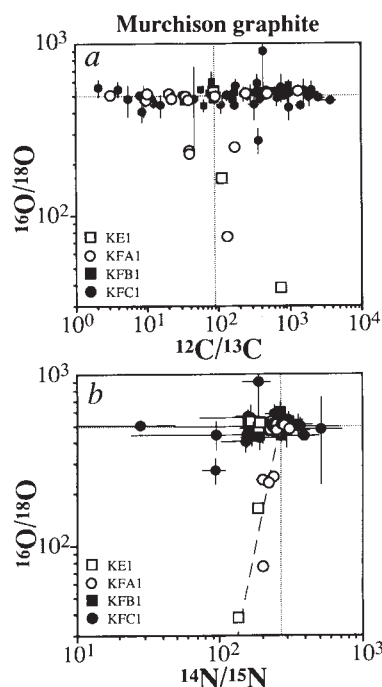


FIG. 4 Isotope ratios in graphite grains; the $^{16}\text{O}/^{18}\text{O}$ ratio is plotted against the $^{12}\text{C}/^{13}\text{C}$ ratio (a) and the $^{14}\text{N}/^{15}\text{N}$ ratio (b). A small fraction of mostly low density (KE1 and KFA1) graphite grains show large ^{18}O excesses. For the KE1 and KFA1 grains these excesses are correlated with ^{15}N excesses (b). Although both Wolf-Rayet stars and the He-burning shells of pre-supernova stars can account for ^{18}O excesses, the former are favoured because 'large' (several micrometres) graphite grains are unlikely to condense in supernova ejecta²³.

of 0.1 during the WN–WC transition, in good agreement with the maximum ratios found in graphite grains (Fig. 3). In pre-supernova stars ^{26}Al is low in the He-burning shell, and mixing with the H-burning shell (where the $^{26}\text{Al}/^{27}\text{Al}$ ratio is ~ 0.2 (ref. 45)) would be required. On the basis of the isotopic signatures alone it is not possible at present to distinguish between Wolf-Rayet stars and pre-supernova stars as the source of ^{12}C -rich and ^{18}O -rich grains. It is, however, doubtful that micrometre-sized grains can grow in supernova ejecta²³.

Finally, the Ne isotopic composition of graphite⁷ cannot be explained by any of the stellar sources considered so far. Helium burning produces high abundances of ^{22}Ne as all ^{14}N is finally turned into ^{22}Ne . But whereas Wolf-Rayet stars in their WC

phase and supernovae are prolific sources of ^{22}Ne , predicted $^{20}\text{Ne}/^{22}\text{Ne}$ ratios are much larger than those observed in graphite (down to 0.0086 ± 0.0021 in KE1; see also Jungck⁴⁷). These predicted ratios are not very sensitive to details of nucleosynthetic models. During H-burning, most of the C and O is converted into ^{14}N . If all this ^{14}N is converted to ^{22}Ne during He-burning, a lower limit on the $^{20}\text{Ne}/^{22}\text{Ne}$ ratio is given by the original $^{20}\text{Ne}/\text{CNO}$ ratio. For solar abundances this ratio is 0.087. Thus graphite grains come either from peculiar AGB or Wolf-Rayet stars with much smaller than solar initial Ne/CNO ratios or, much more likely, some of them must come from novae and their Ne isotopic composition originates from the decay of ^{22}Na (ref. 48). If all of them are as Ne-rich as one graphite spherule analysed by Nichols *et al.*⁴⁹, a fraction of $<1\%$ of such grains is sufficient to account for all the Ne-E(L) in interstellar graphite. The only problem is that the $^{14}\text{N}/^{15}\text{N}$ ratios in all graphite spherules, and the $^{12}\text{C}/^{13}\text{C}$ ratios in most of them, are much larger than those predicted for novae^{50,51}. The question is whether novae with ^{12}C -rich cores can yield isotopically light C early in the explosion, before much ^{13}C is produced (J. W. Truran, personal communication).

We conclude that at least three types of stellar sources, namely AGB stars, Wolf-Rayet stars and novae, are required to account for all the observed isotopic data. □

Received 8 April; accepted 22 September 1993.

- Ott, U. *Nature* **364**, 25–33 (1993).
- Anders, E. & Zinner, E. *Meteoritics* **28**, 490–514 (1993).
- Lambert, D. L. in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 299–316 (Kluwer, Dordrecht, 1991).
- Lewis, R. S., Tang, M., Wacker, J. F., Anders, E. & Steel, E. *Nature* **326**, 160–162 (1987).
- Bernatowicz, T. *et al. Nature* **330**, 728–730 (1987).
- Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **52**, 1235–1244 (1988).
- Amari, S., Anders, E., Virag, A. & Zinner, E. *Nature* **345**, 238–240 (1990).
- Bernatowicz, T. J., Amari, S., Zinner, E. K. & Lewis, R. S. *Astrophys. J.* **373**, L73–L76 (1991).
- Huss, G. R., Hutcheon, I. D., Wasserburg, G. J. & Stone, J. *Lunar planet. Sci.* **23**, 563–564 (1992).
- Nittler, L. R., Walker, R. M., Zinner, E., Hoppe, P. & Lewis, R. S. *Lunar planet. Sci.* **24**, 1087–1088 (1993).
- Lewis, R. S., Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **53**, 3273–3290 (1989).
- Virag, A. *et al. Geochim. cosmochim. Acta* **56**, 1715–1733 (1992).
- Ireland, T. R., Zinner, E. K. & Amari, S. *Astrophys. J.* **376**, L53–L56 (1991).
- Zinner, E., Amari, S., Anders, E. & Lewis, R. S. *Nature* **349**, 51–54 (1991).
- Zinner, E., Amari, S. & Lewis, R. S. *Astrophys. J.* **382**, L47–L50 (1991).
- Ott, U. & Begemann, F. *Astrophys. J.* **353**, L57–L60 (1990).
- Ott, U., Richter, S. & Begemann, F. in *Origin and Evolution of the Elements, Reeves Symp.* (eds Prantzos, N., Vengioni-Flam, E. & Cassé, M.) 528–530 (Cambridge Univ. Press, 1993).
- Prombo, C. A., Podosek, F. A., Amari, S. & Lewis, R. S. *Astrophys. J.* **410**, 393–399 (1993).
- Podosek, F. A., Prombo, C. A., Amari, S. & Lewis, R. S. *Astrophys. J.* (submitted).
- Lewis, R. S., Amari, S. & Anders, E. *Nature* **348**, 293–298 (1990).
- Lewis, R. S., Amari, S. & Anders, E. *Geochim. cosmochim. Acta* (in the press).
- Amari, S., Lewis, R. S. & Anders, E. *Geochim. cosmochim. Acta* (in the press).
- Lattimer, J. M., Schramm, D. N. & Grossman, L. *Astrophys. J.* **219**, 230–249 (1978).
- Lodders, K. & Fegley, B. Jr *Meteoritics* **27**, 250–251 (1992).
- Sharp, C. M. & Wasserburg, G. J. *Lunar planet. Sci.* **24**, 1281–1282 (1993).
- Eberhardt, P., Jungck, M. H. A., Meier, F. O. & Niederer, F. R. *Geochim. cosmochim. Acta* **45**, 1515–1528 (1981).
- Amari, S., Lewis, R. S. & Anders, E. *Lunar planet. Sci.* **21**, 19–20 (1990).
- Lewis, R. S. & Amari, S. *Lunar planet. Sci.* **23**, 775–776 (1992).
- Zinner, E., Wopenka, B., Amari, S. & Anders, E. *Lunar planet. Sci.* **21**, 1379–1380 (1990).

- Virag, A., Zinner, E., Amari, S. & Anders, E. *Geochim. cosmochim. Acta* **55**, 2045–2062 (1991).
- Hoppe, P., Amari, S., Zinner, E., Ireland, T. & Lewis, R. S. *Astrophys. J.* (in the press).
- Ash, R. D., Morse, A. D. & Pillinger, C. T. *Meteoritics* **28**, 318–319 (1993).
- Amari, S., Hoppe, P., Zinner, E. & Lewis, R. S. *Astrophys. J.* **394**, L43–L46 (1992).
- Hoppe, P., Amari, S., Zinner, E. & Lewis, R. S. *Meteoritics* **27**, 235 (1992).
- Gehrz, R. D. in *Interstellar Dust* (eds Allamandola, L. J. & Tielens, A. G. G.) 445–453 (Kluwer, Dordrecht, 1989).
- Lambert, D. L., Gustafsson, B., Eriksson, K. & Hinkle, K. H. *Astrophys. J. Suppl. Ser.* **62**, 373–425 (1986).
- Smith, V. V. & Lambert, D. L. *Astrophys. J.* **294**, 326–328 (1985).
- Dominy, J. F. & Wallerstein, G. *Astrophys. J.* **317**, 810–818 (1987).
- Gallino, R., Busso, M., Picchio, G., Raiteri, C. M. & Renzini, A. *Astrophys. J.* **334**, L45–L49 (1988).
- Busso, M. & Gallino, R. *Astr. Astrophys.* **151**, 205–214 (1985).
- Prantzos, N., Hashimoto, M. & Nomoto, K. *Astr. Astrophys.* **234**, 211–229 (1990).
- Prantzos, N., Doom, C., Arnould, M. & de Loore, C. *Astrophys. J.* **304**, 695–712 (1986).
- Maeder, A. *Astr. Astrophys.* **173**, 247–262 (1987).
- Hashimoto, M., Nomoto, K., Tsujimoto, T. & Thielemann, F.-K. in *Nuclei in the Cosmos* (eds Käppeler, F. & Wisshak, K.) 587–594 (Inst. Phys. Publ., Bristol and Philadelphia, 1993).
- Weaver, T. A. & Woosley, S. E. *Phys. Rep.* **227**, 65–96 (1993).
- Caughlan, G. R. & Fowler, W. A. *Atomic Data and Nuclear Data Tables* **40**, 283–344 (1988).
- Jungck, M. *Pure ^{22}Ne in the Meteorite Orgueil* (Reinhardt, München, 1992).
- Clayton, D. D. *Astrophys. J.* **199**, 765–769 (1975).
- Nichols, R. H. Jr, Hohenberg, C. M., Hoppe, P., Amari, S. & Lewis, R. S. *Lunar planet. Sci.* **23**, 989–990 (1992).
- Wiescher, M., Görres, J., Thielemann, F.-K. & Ritter, H. in *Nucleosynthesis and its Implications on Nuclear and Particle Physics* (eds Audouze, J. & Mathieu, N.) 105–112 (Reidel, Dordrecht, 1986).
- Woosley, S. E. in *Nucleosynthesis and Chemical Evolution* (eds Hauck, B., Maeder, G. & Meynet, G.) 1–95 (Observatoire de Genève, 1986).

ACKNOWLEDGEMENTS. S.A., P.H. and E.Z. thank R. Walker for his interest and support. We gratefully acknowledge the role of E. Anders in the isolation of interstellar graphite. This work was funded by NASA (S.A., R.L. and E.Z.), the US National Science Foundation (E.Z.) and the Schweizerische Nationalfonds zur Förderung der wissenschaftlichen Forschung (P.H.).

The structure of crystalline profilin- β -actin

Clarence E. Schutt*, James C. Myslik*, Michael D. Rozycki*,
Nalin C. W. Goonesekere* & Uno Lindberg†

* Department of Chemistry, Henry H. Hoyt Laboratory, Princeton University, Princeton, New Jersey 08544, USA

† Department of Zoological Cell Biology, WGI, Arrhenius Laboratories for Natural Sciences, Stockholm University, S-10691 Stockholm, Sweden

The three-dimensional structure of bovine profilin- β -actin has been solved to 2.55 Å resolution by X-ray crystallography. There are several significant local changes in the structure of β -actin compared with α -actin as well as an overall 5° rotation between its two major domains. Actin molecules in the crystal are organized into ribbons through intermolecular contacts like those found in oligomeric protein assemblies. Profilin forms two extensive contacts with the actin ribbon, one of which appears to correspond to the solution contact *in vitro*.

THE discovery of myosin and filamentous actin (F-actin) in non-muscle cells intensified the search for mechanisms underlying force generation and movement in eukaryotic systems. This search led to the identification of proteins that interact with actin to modify its behaviour, forming a basis for the repertoire of structures and movements observed in non-muscle cells^{1,2}. The presence of non-filamentous forms of actin in extracts of various non-muscle cells indicated that the control of filament assembly and disassembly may be a feature of the dynamic behaviour of cells^{3,4}.

The first hint that actin-binding proteins are important in controlling filament assembly came from the observation that binding of monomeric actin to DNase I inhibits both DNA hydrolysis and F-actin assembly⁵. Subsequently, the crystalline form of the DNase I inhibitor⁶ was found to consist of non-muscle actin and a second actin-binding protein, profilin, in a 1:1 ratio^{7,8}. Profilin inhibited the assembly of F-actin *in vitro* by sequestering monomeric actin⁸, and evidence followed that the profilin-actin heterodimer is a cellular storage form of monomeric actin⁹⁻¹¹. Thymosin β_4 is another component that controls F-actin assembly¹².

Profilin seems to link receptor-mediated transmembrane signalling and microfilament-based motility¹³. Hydrolysis of phosphatidylinositol (4,5)-bisphosphate (PtdInsP₂) by phospholipase C_γ1 is inhibited by profilin, unless the enzyme has been activated by receptor kinase-dependent tyrosine phosphorylation^{14,15}. Specific binding of PtdInsP₂ to profilin dissociates the profilin-actin complex^{13,16,17}. Thus PtdInsP₂, appearing in the inner leaflet of the lipid bilayer of the plasma membrane as a result of transmembrane signalling, recruits profilin-actin to the site of filament assembly^{13,17}.

We report here the crystallographic determination (Table 1) of the bovine profilin- β -actin¹⁸ structure at 2.55 Å resolution. X-ray amplitudes were phased with a combination of multiple isomorphous replacement and molecular replacement using the structure of α -actin¹⁹ as a search model. This structure confirms the interpretation at low resolution that profilin- β -actin assembles into a higher-order structure in which actin molecules make extensive contacts across a 2₁ screw axis and profilin molecules intercalate between consecutive actin molecules lying along each edge of the actin 'ribbon'²⁰. The actin-actin ribbon contact is more characteristic of an oligomeric protein contact than an incidental crystal contact. Considering that the profilin- β -actin crystal can be converted to a well diffracting, semicrystalline fibre under conditions known to dissociate the profilin-actin complex²⁰, the ribbon contact may be related to actin-actin contacts in F-actin, providing a starting point for constructing a model of the β -actin filament.

Comparison of β - and α -actins

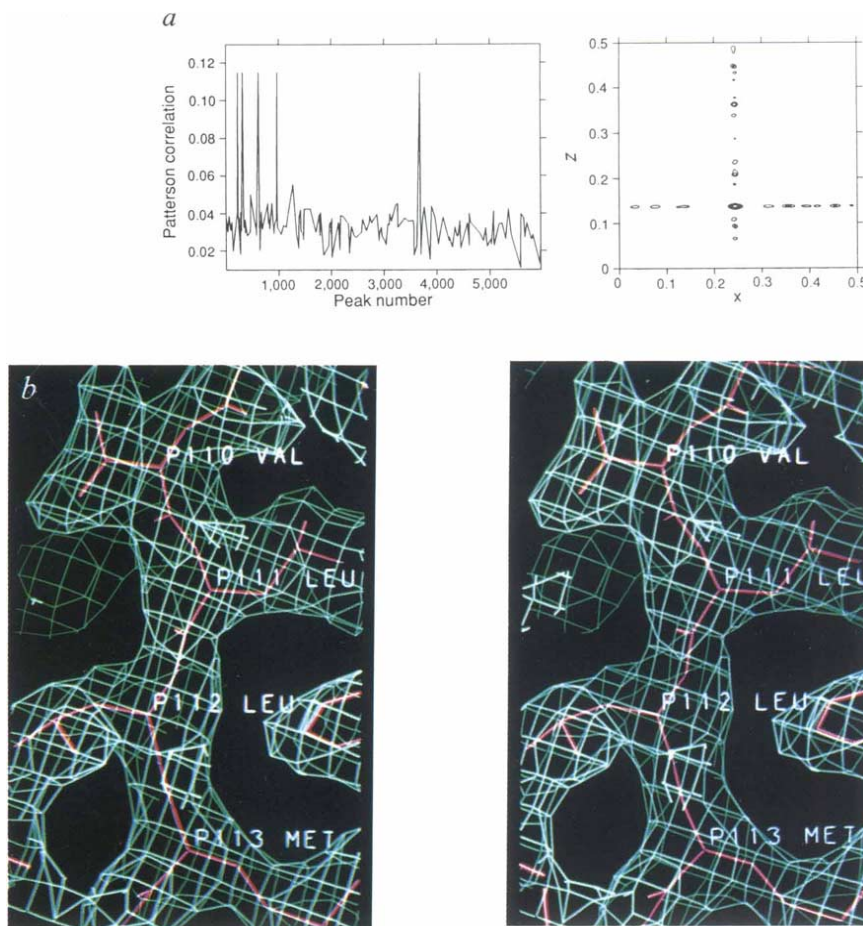
The α -actin molecule is composed of two major domains, the

smaller divided into subdomains 1 and 2, the larger divided into subdomains 3 and 4 (ref. 19). Bovine β -actin is generally similar in structure to α -actin, but a number of significant differences merit explanation. Alignment of α -carbon atoms in subdomain 3 (Fig. 2a), leads to superposition of a major portion of β -actin subdomain 4 onto α -actin, although local displacements of residues are observed at the top of subdomain 4 (residues 192-208 and 230-249, using the β -actin sequence of ref. 21) and the bottom of subdomain 3 (residues 275-291). A rotation of the small domain of β -actin relative to α -actin becomes apparent after this alignment. Most of subdomains 1 and 2 of β -actin can be superimposed onto α -actin after a torsion change of ($\Delta\phi = +5^\circ$, $\Delta\psi = +2^\circ$) at Ala 138, although residues 338-372 of β -actin are not moved in this operation and require a separate torsion change of ($\Delta\phi = -5^\circ$, $\Delta\psi = +2^\circ$) at Ser 338. Several stretches of residues are not superimposable by these torsion changes, including 2-5, 39-50 and 55-65. Normal mode analysis of monomeric α -actin predicted a similar hinging between the two domains in α -actin²². Hexokinase has been observed in two conformations related by rotations in a homologous region²³.

Differences in amino-acid conformations between the two actin isoforms fall into four categories. The first and second groups include residues whose main-chain atoms can be superimposed by one of the operations described, and whose resulting side-chain conformations are very similar, with only subtle differences in their interactions with neighbouring residues. Residues in the first group have side chains buried away from solvent, whereas those in the second group are partially or fully solvent-accessible. Residues that bind ATP fall into these two categories. As a result, the Sr²⁺-ATP complex in β -actin lies in an environment similar to the Ca²⁺-ATP complex in α -actin, and the cation-nucleotide complex itself adopts an almost identical conformation in the two isoforms. The third group consists of residues with significant changes in side-chain conformations. Residues in this group may project freely into solvent without interacting with other residues, or they accommodate different intermolecular contacts in the two crystals. For example, the phenol group of Tyr 169 in β -actin is rotated by 140° relative to its position in α -actin, allowing it to contact the side chain of Ile 73 on neighbouring profilin. The fourth group includes residues whose main-chain conformations are not superimposable by the operations described, leading to significantly changed side-chain orientations. Among the α/β sequence differences, residues Cys/Val 10, Val/Cys 17, Thr/Val 103, Val/Thr 129, Leu/Met 153, Asn/Thr 163, Thr/Ala 260, Asn/Thr 297 and Met/Leu 299 fall into the first group; Leu/Met 16, Ile/Val 76, Met/Leu 176, Asn/Thr 201, Ile/Leu 267, Ala/Cys 272, Tyr/Phe 279, Ile/Val 287, Thr/Ser 358 and Ala/Ser 365 are in the second group; Asn/Gln 225 is in the third group; and Glu/Asp 2, Glu/Asp 4, Thr/Ile 5 and Thr/Ala 6 fall in the fourth.

Several differences between the actin isoforms involve coordi-

FIG. 1 Structure determination of bovine profilin- β -actin. **a**, Molecular replacement using a search model described in Table 1 legend. The search model was rotated and translated as a single, rigid unit using molecular replacement as implemented in X-PLOR⁴⁶. The left-hand panel shows the refined Patterson correlation coefficients for the highest 6,000 rotation function solutions in the resolution range from 8.00 to 4.00 Å. The six highest peaks, one of which is occluded by the vertical axis, have Patterson correlation coefficients of 0.115 (4 σ) and represent an equivalent unique solution at the pseudo-orthogonal eulerian angles of (75.2°, 72.9°, 84.4°). The right-hand panel gives the translation function calculated for the solution to the rotation function in the same resolution range. The plot is contoured from 5 σ to maximum by σ , with the highest peak ($T=0.332$, 12 σ) at ($x=0.219$, $y=0.247$, $z=0.140$). A full description of this procedure is presented in ref. 50. **b**, Portion of atomic model of profilin residues 110–113 (pink). The accompanying $2F_o - F_c$ omit map (green cage) was calculated after removal of an 8 Å radius sphere of the partially refined profilin- β -actin-ATP-Sr²⁺ model (Table 1 legend), centred at the α -carbon of profilin residue Leu 111, followed by simulated annealing refinement of the remainder of the model using X-PLOR. The map contouring level is 0.50 σ .



nated changes in groups of residues. The first of these includes residues 37 and 52, which bind DNase I (ref. 19). Residues 37, 38 and 52 are nearly superimposable in the two isoforms, but residues deviate increasingly towards the middle of the loop. The side chains of Arg 37, Arg 39, Val 43, Met 44, Met 47 and Asp 51 show significant differences in orientation, which may be explained by the different crystalline environments of the isoforms. In α -actin, Gly 42, Val 43 and Met 44 form a β -strand that is incorporated into a β -sheet in DNase I (ref. 19), whereas in profilin- β -actin these residues form an actin-actin contact. Aspartic acid at position 51 forms a salt bridge with Arg 37 in β -actin which is not found in α -actin. This region appears to be one of relative mobility in β -actin, as evidenced by an average atomic temperature factor of $35 \pm 3 \text{ \AA}^2$ for these residues, compared with $20 \text{ \AA}^2$ for the entire profilin-actin structure.

Structural differences in the amino-terminal region of actin are important because of its role in binding myosin and other actin-binding proteins²⁴. Five α/β differences lie in the span of residues 1–6, including Asp 1, which is not present in β -actin²¹. This stretch is helical in α -actin. Substitution of Thr by Ile at residue 5 brings this residue closer to His 101 and Pro 102 in β -actin, causing a displacement relative to α -actin of Asp 2 through to Asp 4 (Fig. 2b) and a conformation change to a non-helical turn for these residues. Atomic temperature factors for residues 2–5 of β -actin average $33 \pm 5 \text{ \AA}^2$, indicating a relatively high positional uncertainty as found in α -actin¹⁹.

Residues 373–375 of α -actin were removed before crystallization with DNase I (ref. 19). The local environment of these residues in crystalline profilin- β -actin is shown in Fig. 2c. The terminal α -helix starting at Ile 369 continues to Cys 374 in the β -actin structure with normal helical geometry. Lysine 373 extends underneath the helix to interact with Glu 361, while Cys 374 tucks into a pocket containing Tyr 133, Val 134 and Ile 357. The aromatic ring of Phe 375 is surrounded by hydrophobic side

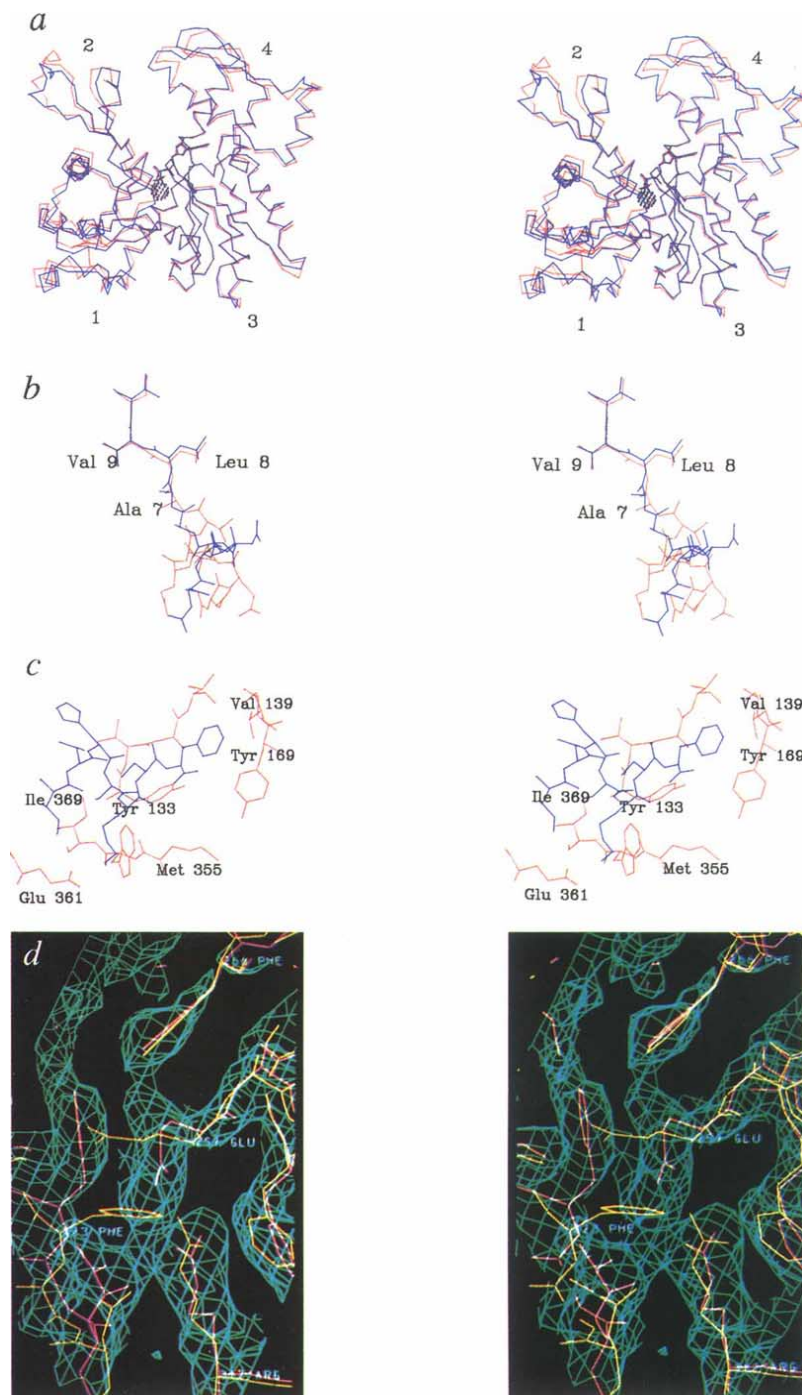
chains from Tyr 133, Ile 136, Val 139, Tyr 169, Ala 170 and Met 355. The carboxy terminus of β -actin has a stability comparable to the entire profilin-actin complex, as evidenced by average atomic temperature factors of $22 \pm 5 \text{ \AA}^2$ for residues 351–375, which may be due either to its interaction with profilin, described later, or to the presence of those residues missing in the α -actin crystal structure. Temperature factors above $40 \text{ \AA}^2$ have been reported in this region for α -actin¹⁹.

A difference between the two actin isoforms arises in the contact between the main body of subdomain 4 and the loop consisting of residues 264–273. In α -actin, the aromatic ring of Phe 266 lies in an energetically unfavourable environment next to the carboxylate group of Glu 259 (Fig. 2d). This instability has led Holmes *et al.* to propose that residues in this loop spontaneously refold in solution, forming a 'hydrophobic plug' which provides essential stabilization energy for their filament model²⁵. In β -actin, the side-chain conformations of Phe 223 and Glu 259 are different, allowing for favourable interactions with the phenyl group of Phe 266 as well as a new electrostatic interaction of the carboxylate group of Glu 259 with Arg 312. This arrangement appears to stabilize the conformation of the 264–273 loop as it is found in the crystal. The average atomic temperature factor of $22 \pm 5 \text{ \AA}^2$ for residues 264–273 indicates relative stability for this region which may, however, derive from its proximity to a contact with profilin. Although three α/β sequence differences fall within this region (Ala 260, Leu 267 and Cys 272), they do not account for the differences because the two structures are superimposable in this region apart from the changes to Phe 223 and Glu 259 already described.

Profilin structure

Profilin is bisected by an antiparallel β -pleated sheet, as shown in Fig. 3a. Both termini are α -helical and pack against the same side of the central sheet, connecting to it by short loops. The

FIG. 2 Comparison of α - and β -actin. *a*, Comparison of α -carbon paths of β -actin (blue) and α -actin (red). The locations of the 4 subdomains are labelled according to ref. 19. Residues in subdomain 3 of the two isoforms were superimposed using a least-squares alignment of α -carbons for residues 135–182 and 263–335 using the program QUANTA (Polygen). The r.m.s. positional deviation of α -carbons for these residues was 0.497 Å. *b*, Comparison of the amino-terminal residues of β -actin (blue) and α -actin (red). Residues Ala 7, Leu 8, and Val 9 are labelled. The remaining residues shown in α -actin are Asp 1, Glu 2, Asp 3, Glu 4, Thr 5, Thr 6 and Cys 10, and in β -actin, Asp 2, Asp 3, Asp 4, Ile 5, Ala 6 and Val 10. *c*, Carboxy terminus of β -actin. Residues 369–375 are shown in blue, residues Tyr 133, Val 134, Ala 135, Ile 136, Val 139, Tyr 169, Ala 170, Met 355, Trp 356, Ile 357 and Glu 361 are shown in red. The amino-terminal residue is labelled in each fragment. *d*, Comparison of side-chain orientations of Phe 223, Asp 259, Phe 266 and Arg 312 in β -actin (pink) and α -actin (yellow). Accompanying $2F_o - F_c$ omit map (green cage) was calculated after omission of an 9 Å-radius model sphere, centred at the α -carbon of β -actin residue Phe 223, and simulated annealing refinement of the remainder of the model using X-PLOR (see Table 1 legend). The map contouring level is 0.50 σ .



amino-terminal helix (H1) packs flat against strands 1, 2 and 7, and the carboxy-terminal helix (H4) lies next to it against strands 5, 6 and 7. The two helices do not pack tightly against each other, but contribute side chains to a common hydrophobic core involving the β -sheet, as well as to a salt bridge between Asp 8 and Lys 126. Residues on the opposing face of the central sheet contribute to a second, more extensive hydrophobic core populated by aliphatic residues and bounded by two small α -helices (H2 and H3) and a short β -hairpin containing strands 3 and 4. This latter core is sealed from above by the loop connecting strands 1 and 2, and at the side by a helical turn preceding helix 2. A β -bulge at Ile 73 redirects strand 4, bending it into the central sheet. This bend is stabilized by an interaction between Arg 55 and Asp 75. Strands 5 and 6 are connected by a type-II turn and form a protrusion from the profilin surface. The

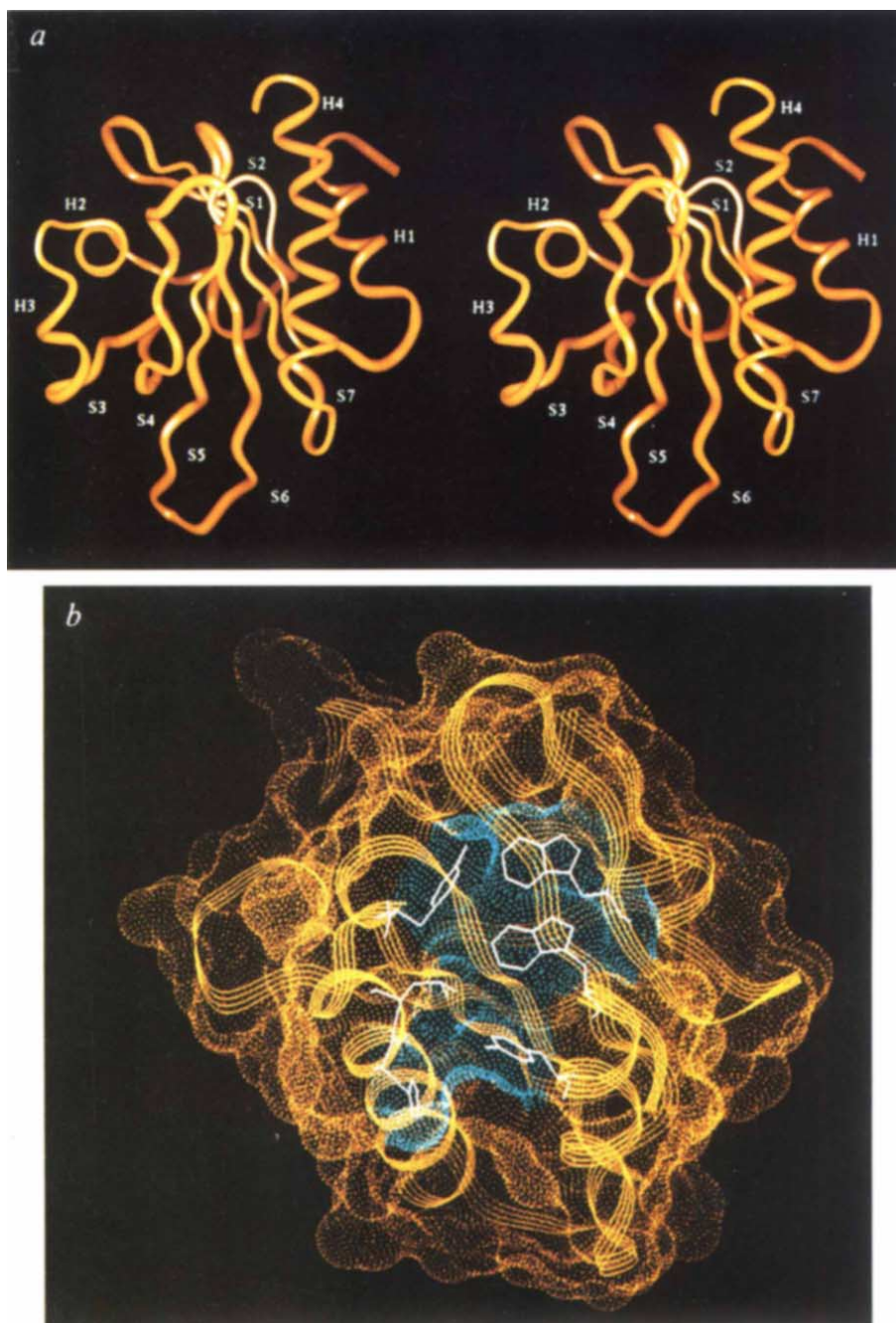
polypeptide fold of strands 3 to 6 of bovine profilin is similar to those found for several SH3 domains²⁶.

Alignment of 16 profilin sequences in the Swiss Protein Database showed that residues Trp 3, Tyr 6, Trp 31 and Leu 134 are strictly invariant, except in the putative profilin sequence from the virus vaccinia (Table 2); His 133 and Tyr 139 are also highly conserved in these sequences. These residues lie on a solvent-exposed surface between helix 1 and helix 4 (Fig. 3*b*). A second extensive, but less strongly conserved, hydrophobic surface patch contains residues Tyr 24, Pro 28, Pro 44, Val 51, Leu 78, Phe 83, Met 104, Ala 106, Tyr 128 and Arg 135.

Protein-protein interactions

The extensive crystal contact between α -actin and DNase I enabled a correspondence between the crystalline and solution

FIG. 3 Structure of bovine profilin. **a**, Polypeptide fold of profilin. Labelled elements of secondary structure include helix 1 (H1), residues 2–11; helix 2 (H2), 44–51; helix 3 (H3), 57–61; helix 4 (H4), 120–136; strand 1 (S1), 17–24; strand 2 (S2), 32–34; strand 3 (S3), 62–64; strand 4 (S4), 69–76; strand 5 (S5), 83–90; strand 6 (S6), 97–104 and strand 7 (S7), 109–114. **b**, Highly conserved hydrophobic residues in profilin form a solvent-exposed hydrophobic surface. The polypeptide fold of profilin is depicted as a yellow ribbon. Residues in white are (clockwise from top right) Trp 31, Trp 3, Tyr 6, His 133, Leu 134 and Tyr 139. Solvent-exposed surfaces are depicted for these residues by blue–green dots and for the remainder of profilin in brown dots. The probe radius for solvent-exposed surfaces was 1.40 Å.



binding sites of DNase I and α -actin to be established¹⁹. We expected a similar result for crystalline profilin- β -actin and considered it likely that profilin, in view of its proposed function as a monomer-sequestering protein, would bind to β -actin to block the actin-actin contacts in filaments. But analysis of low-resolution electron density maps revealed that both profilin- β -actin and β -actin- β -actin contacts are extensive in the crystalline state²⁰. The present high-resolution structure confirms this view by verifying the molecular boundaries and enabling intermolecular contacts to be evaluated.

The high-resolution structure of profilin- β -actin is presented in Fig. 4a. Profilin forms two major contacts with actin in the crystal. The primary contact comprises a region of profilin defined by helix 3, the amino-terminal portion of helix 4, and strands 4, 5 and 6. This region makes contact with a site on actin spanning the large and the small domains of the molecule at the bases of subdomains 1 and 3. Formation of this contact buries a combined interfacial surface area on profilin and actin

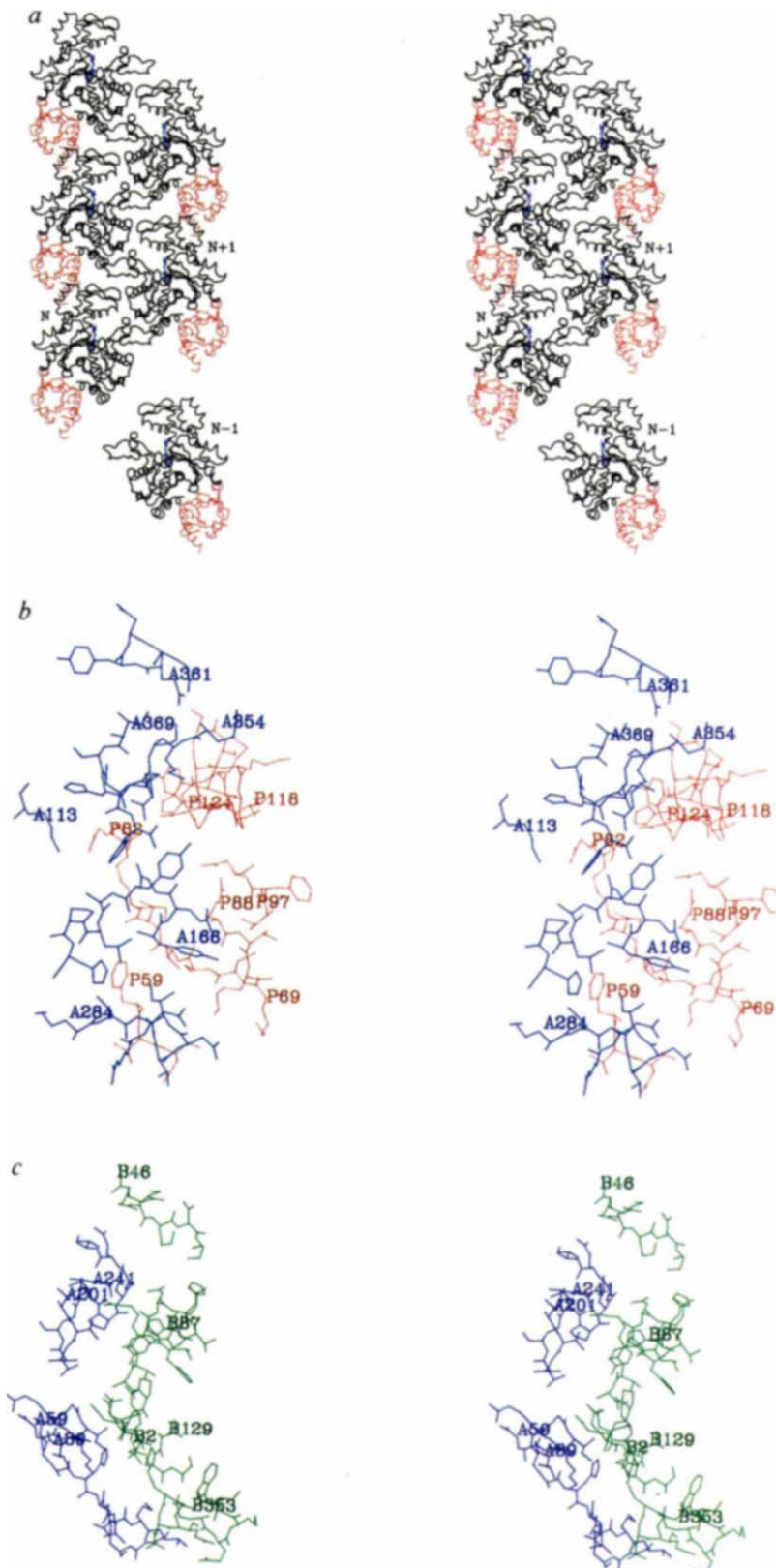
of 2,250 Å². Twenty-one residues each from actin and profilin contribute to this contact (Fig. 4b), resulting in a close-packed, complementary surface of ionic, polar and hydrophobic interactions, including ten hydrogen bonds. The buried surface area of this contact is approximately ten times that of a typical protein:protein crystal contact and lies in the range observed for functionally important protein interfaces such as those found in protease-inhibitor and antibody-antigen complexes²⁷. For comparison, the crystal contact which has been identified as the solution binding site between DNase I and α -actin¹⁹ buries a combined area of 1,830 Å² between the two proteins.

Biochemical studies indicate that the primary profilin-actin contact corresponds to the contact observed in solution. Removal of Phe 375 from β -actin by carboxypeptidase A digestion decreases the stability of the profilin-actin complex²⁸. This finding is consistent with the central role of Phe 375 in this contact; the phenyl ring interacts with Ile 73, His 119, Gly 121 and Asn 124 of profilin, while the terminal carboxylate group forms

a salt bridge with Arg 74 of profilin. A second observation is that Lys 125 of profilin and Glu 364 of actin form an intermolecular electrostatic interaction at the profilin-actin contact site. Glutamate 364 of actin has been chemically crosslinked to Lys 115 of profilin from *Acanthamoeba castellanii*²⁹, which in sequence alignment is homologous to Lys 125 of bovine profilin. Finally, profilin-actin dissociates when the pH is less than 6.0 (ref. 30) suggesting that histidyl residues are involved in the

intermolecular contact. The profilin-actin contact at subdomains 1 and 3 of actin is the only contact involving histidyl residues; the imidazole ring His 173 of actin stacks upon the phenyl ring of Phe 59 of profilin and His 119 of profilin sits in a pocket formed by Tyr 169, Met 355 and Phe 375 of actin. Protonation of these residues leading to formation of positively charged imidazolium ions would destabilize these apolar interactions.

FIG. 4 Organization of profilin and actin in profilin- β -actin crystals. *a*, The profilin-actin ribbon. Actin and profilin are depicted in black and red, respectively. ATP is shown as a blue stick figure, and Sr^{2+} as a blue sphere at the van der Waals radius. Five profilin-actin heterodimers are shown in an intact ribbon. The n th monomer lying at the bottom of the ribbon makes an extensive contact with monomer $n+1$ lying across a 2_1 screw axis lying in the plane of the page parallel to the long axis of the ribbon. A sixth heterodimer ($n-1$) is shown removed from the ribbon to show the solution complex. *b*, The primary profilin-actin contact. Amino-terminal residues only are labelled in each fragment except for P124 which has been labelled for orientation. Actin and profilin are in blue and red, respectively. The contact can be divided into two parts. The first involves contacts between Glu 82, Val 118, His 119, Gly 121, Met 122, Asn 124, Lys 125, Tyr 128 and Glu 129 of profilin and residues of β -actin subdomain 1 including Lys 113, Gln 354, Met 355, Glu 361, Glu 364, Ile 369, His 371, Arg 372, Lys 373 and Phe 375. The second part includes profilin residues Phe 59, Val 60, Asn 61, Lys 69, Ser 71, Val 72, Ile 73, Arg 74, Arg 88, Lys 90, Thr 97 and Asn 99, and β -actin subdomain 3 residues Tyr 166, Glu 167, Tyr 169, Leu 171, Pro 172, His 173, Lys 284, Asp 286, Val 287, Asp 288 and Arg 290. Total buried surface area was calculated according to ref. 51 using X-PLOR. The probe radius was 1.4 Å. A residue was considered to participate in a contact if its solvent-accessible surface area in the isolated monomer decreased by 10 Å² or more in forming the interaction. The criteria for hydrogen bonding were (1) a distance between donor and acceptor atoms no greater than 3.4 Å, and (2) an angle of 110° or more between donor atom, acceptor atom and antecedent. *c*, The actin-actin ribbon contact. Only amino-terminal residues are labelled in each fragment. The actin molecule drawn in green (molecule B) is related to the one drawn in blue (molecule A) by the operation of a 2_1 -screw axis parallel to the crystallographic b -axis. The interface can be dissected into two subcontacts. At the bottom of the figure, the small domain of molecule A (His 40, Gln 41, Gly 42, Met 44, Val 45, Gly 46, Ser 60, Lys 61, Gly 63 and Ile 64) makes contact with the small domain of molecule B (acetyl-Asp 2, Asp 4, Asp 5, Glu 99, Glu 100, Pro 130, Ala 131, Gln 353, Ser 358 and Gln 360). At the top, the large domain of molecule A (Thr 202, Ala 204, Gly 205, Pro 243, Asp 244, Gly 245, Gln 246 and Val 247) makes contact with the small domain of molecule B (Gly 48, Lys 50, His 87, Tyr 91, Asn 92, Arg 95 and Ala 97). Both of the subcontacts are predominantly apolar except for the side chains of His 40 and Asp 2, and Glu 205 and Arg 95. Total buried surface area was calculated as in *b*.



Each actin molecule in crystalline profilin- β -actin makes contact with eight other actin molecules²⁰. The most extensive contact is made between two actin molecules related to one another by a 2₁ screw axis parallel to the crystallographic *b*-axis, giving rise to the zig-zag ribbon structure shown in Fig. 4a. The interface at this 'ribbon contact' buries 1,670 Å² of total surface area

TABLE 1 Crystallographic analysis

	Native	K ₂ O ₈ O ₄	K ₂ PtCl ₄
Resolution (Å)	2.55	2.55	2.55
Unique reflections	16,158	15,160	15,156
Completeness (%)	98.8	92.7	92.7
R _{sym} (%) (<i>I</i> > 0)	6.5	7.1	7.3
Refinement (8.0–2.55 Å)			
Resolution range (Å)	Reflections/completeness (<i>I</i> > 3σ(<i>I</i>)/%)	R-factor (%)	R _{cum} [†] (%)
4.76–8.00	1,949/94	19.35	19.35
3.92–4.76	1,968/98	15.06	17.07
3.47–3.92	1,946/97	17.65	17.24
3.18–3.47	1,877/98	19.80	17.70
2.96–3.18	1,836/96	21.83	18.21
2.80–2.96	1,732/94	25.82	18.90
2.66–2.80	1,688/83	27.67	19.50
2.55–2.66	1,539/82	29.61	20.05
r.m.s. deviation, bond length	0.018 Å		
r.m.s. deviation, bond angle	3.84°		
Overall restrained B-factor	19.9 Å ²		

Calf spleen profilin- β -actin was purified and crystallized as described^{6,7,18}. X-ray film data was processed as described for native and derivatized crystals⁴⁴. Crystallographic analysis was carried out using the CCP4 program suite⁴⁵. Multiple isomorphous replacement (MIR) phases were obtained to better than 2.6 Å and are described in detail in ref. 20; however, only K₂O₈O₄ and K₂PtCl₄ derivatives were used here. The MIR electron density map contained elements of recognizable secondary structure but sufficient ambiguities remained in connecting regions that sequence assignment could not be completed without additional phase information from molecular replacement (MR). A search model was used containing amino-acid residues 7–38 and 52–348 of α -actin, without nucleotide or divalent cation. Remaining residues were omitted because of high-temperature factors at the termini¹⁹ and because structural changes were anticipated at the DNase I-binding site. The search model was rotated and translated in the profilin- β -actin unit cell using the macromolecular analysis program X-PLOR (version 3.0)⁴⁶. A solution was obtained for the unique alignment of the model with β -actin (Fig. 1a). Rotational and translational searches were made for the four individual subdomains of α -actin, and solution positions were nearly identical to those for the intact molecule. Residues in β -actin which differ from α -actin²¹ were substituted into the MR solution, and Sr²⁺-ATP, used for profilin- β -actin crystallization²⁰, was fit to the electron density. The resulting starting model for β -actin was refined by conjugate-gradient minimization in X-PLOR using native X-ray amplitudes between 10.0 and 3.4 Å. Phases calculated from this refined model and amplitudes from native and OsO₄-derivative data were used to calculate difference Fourier transforms which gave peaks at the previously identified²⁰ OsO₄²⁻ sites in the unit cell. Simulated annealing at this point led to degradation of phase quality observed as increase in noise in the difference Fourier, an effect that was not evident after completion of model building. Model phases were combined with MIR phases and an electron density map calculated. Residues 2–6, 39–40 and 349–375 of β -actin (numbered according to ref. 21) were modelled using FRODO⁴⁷, augmented with the structural database searching program FRAGLE⁴⁸. An initial model for roughly half of the 139 residues of profilin was made by building secondary structural elements and connecting loops as polyalanine fragments. A new phase-combined map was calculated adding the newly modelled fragments of actin and profilin to the first model. This map showed sufficient continuity and side-chain density to model the amino-acid sequence⁴⁹ of bovine profilin. The resulting profilin- β -actin model was completed apart from residues 41 to 51 of β -actin and 138 and 139 of profilin, which were still difficult to fit to electron density. The model was refined by simulated annealing from 3,000 K in X-PLOR between 8.0 and 2.6 Å. Independence of the refined model from the α -actin search model was evaluated by examining each residue for fit to an 'omit map'. Profilin residues (1–137) and residues 6–35, 70–135, 335–375 of actin constituted a half-model which was divided into overlapping spheres of radii 7–9 Å. For each sphere, an annealed omit map⁴⁶ was calculated between 6.00 and 2.60 Å. Residues within each sphere were examined for agreement to electron density in the map and adjusted as necessary to improve their fit. Phases from the adjusted half-model were combined with MIR phases to produce another map. The remaining residues in β -actin were examined in this map, including the Sr²⁺-ATP complex, and the previously unfit residues 41–51 of actin and 138–139 of profilin were inserted. The completed profilin- β -actin-Sr²⁺-ATP model was refined between 8.00 and 2.55 Å in X-PLOR using simulated annealing followed by restrained refinement of atomic temperature factors. Coordinates of the complex will be deposited in the Brookhaven Protein Database.

* $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$, where *I* is the intensity of a given reflection, and $\langle I \rangle$ is the mean intensity of symmetry-related reflections.

† R_{cum} is the cumulative R-factor.

between the two monomers involved. By comparison, the most extensive actin-actin contact in the DNase I- α -actin crystal buries a combined surface area of only 150 Å². Seventeen residues are contributed from each protein at the ribbon contact. No hydrogen bonds are found at this contact. Instead, hydrophobic interactions predominate (Fig. 4c), even for residues containing charged groups. Exceptions to this are the Arg 95:Glu 205 and His 40:Asp 2 interactions. The ribbon contact is comparable in extent to the primary profilin- β -actin and DNase I/ α -actin contacts, and is significantly larger than contacts found typically in crystals. As with the primary profilin-actin contact, it lies in the range observed for known protein interfaces, although the close packing of predominantly hydrophobic groups is more characteristic of oligomeric protein contacts than of protease-inhibitor or antibody-antigen complexes^{27,31}.

The remaining actin-actin contacts in the profilin- β -actin crystal occur through interactions between symmetry-related ribbons. The second largest actin-actin contact in the crystal is 40% as extensive as the ribbon contact. In its interaction with the actin ribbon, profilin forms a contact with a second actin molecule, fitting the amino-terminal helix of profilin into a groove formed by three of the four helices of subdomain 4. Thus profilin appears to act as a wedge between adjacent actin molecules, stabilizing the ATP-containing ribbon in the crystal (Fig. 4a). This second profilin-actin interaction buries 1,110 Å² of combined interfacial area, contributed by 14 residues from actin and 11 from profilin. Only one hydrogen bond is found, and the interactions at this site are mainly hydrophobic. These characteristics are more like the actin-actin ribbon contact than the primary profilin-actin contact.

Implications of the ribbon structure

Actin functions as a self-assembling protein, and the ribbon contact may be related to the actin-actin contacts in actin filaments. It is not unusual for protein crystals and fibres to use similar intermolecular contacts. Crystallization of sickle deoxyhaemoglobin S (HbS) seems to compete with formation of 14-strand fibres³². Within the resolution limits of electron microscopy and fibre diffraction, it appears that both processes incorporate the same basic functional unit, a ribbon of two half-staggered strands of HbS tetramers³². The combined buried interfacial surface area between contacts within HbS ribbons (880 Å² for the lateral, 890 Å² for the axial tetramer contacts) are comparable to the actin-actin contacts in the profilin- β -actin ribbons. In the bacterial RecA protein, crystal contacts have been used to model the inactive form of the RecA filament³³.

On the basis of two observations, we have proposed that the actin ribbon represents a twisted and stretched form of filamentous actin stabilized by profilin²⁰. First, assembly of paracrystalline F-actin competes with crystal growth in solution and, by analogy with HbS, may use closely related actin-actin contacts²⁰. Second, the profilin- β -actin crystal lattice can be converted to a well diffracting, semi-crystalline fibre after transfer to low pH²⁰, which dissociates profilin from actin³⁰. In addition, three-dimensional reconstructions of F-actin imaged by cryo-electron microscopy show a striking resemblance between the wide view of the filament and the profilin-actin ribbon³⁴. The present high-resolution structure extends our previous studies on the actin ribbon by providing evidence for a specific oligomeric interaction. The footprint of one actin molecule on its neighbour in the ribbon includes a region of the sequence that is nearly identical to that associated with the DNase I- α actin contact. DNase I binds both the small and large domains of actin at subdomains 2 (residues 39–46, 60–64) and 4 (residues 202–204 and 207)¹⁹. Thus DNase I, which inhibits filament assembly, contacts all of the residues in subdomain 2 and two of the seven residues in subdomain 4 that are involved in the ribbon contact.

The role of profilin

Compared with α -actin, β -actin as it appears in the profilin-

TABLE 2 Conservation of solvent-exposed hydrophobic residues in profilin

Bovine	Residue (aligned with bovine sequence)					
	W3	Y6	W31	H133	L134	Y139
Human	W	Y	W	H	L	Y
Mouse	W	Y	W	H	L	Y
White birch	W	Y	W	Y	L	L
Sea urchin	W	Y	W	Y	L	M
Sand dollar	W	Y	W	Y	L	M
<i>Drosophila</i>	W	Y	W	Y	L	Y
<i>Physarum</i> I	W	Y	W	Y	L	Y
<i>Physarum</i> II	W	Y	W	Y	L	Y
<i>Tetrahymena</i>	W	Y	W	T	L	Y
<i>Dictyostelium</i> I	W	Y	W	Y	L	Y
<i>Dictyostelium</i> II	W	Y	W	Y	L	C
<i>Acanthamoeba</i> I	W	Y	W	Y	L	F
<i>Acanthamoeba</i> II	W	Y	W	Y	L	F
<i>Saccharomyces</i>	W	Y	W	Y	L	Y
Vaccinia virus	W	I	L	R	V	N

actin ribbon has a 5° rotation between its two major domains. This rotation could be invoked to account for the increased rate of nucleotide exchange on actin in the presence of profilin³⁵. However, the structural similarity between the nucleotide-binding sites in the two actin isoforms argues against this explanation. Furthermore, it is possible that the domain rotation represents the native structure of β -actin or a conformation induced during formation of the ribbon contact.

It is surprising that neither of the two contacts between profilin and actin use the conserved hydrophobic surfaces of profilin. The more highly conserved surface shown in Fig. 3b lies between the two terminal helices that bind actin, suggesting that these residues have a regulatory role. The specific affinity of profilin for poly(L-proline)³⁶ is abolished by mutagenesis of two of the

residues in this region (Trp 3 or His 133)³⁷. SH3 domains bind proline-rich sequences downstream of receptor-linked tyrosine kinases³⁸, and the similarities between bovine profilin and SH3 domains suggest that they carry out related docking functions in signalling pathways. The ability of profilin to bind simultaneously to actin ribbons and proline-rich sequences could serve an important function in actin-based motility. For example, this could explain the role of profilin in establishing links between growing actin filaments and the bacterial protein Act A at the tail of *Listeria monocytogenes*³⁹, providing propulsion for bacterial cells through the eukaryotic cytoplasm⁴⁰.

A second role for the conserved hydrophobic surface shown in Fig. 3b is suggested by the fluorescence quenching of Trp 3 and Trp 31 in the presence of PtdInsP₂ (ref. 41). These residues lie adjacent to a proposed PtdInsP₂ binding sequence⁴² on helix 4. If fatty acid chains of PtdInsP₂ bind the hydrophobic surface, the four positively charged residues in helix 4 (Lys 125, Lys 126, Arg 135 and Arg 136) could bind the phosphoinositol head group, accounting for the ability of PtdInsP₂ to dissociate bovine profilin-actin *in vitro*¹³. But as micellar PtdInsP₂ binding to profilin was used, further investigation is necessary to determine whether a single-molecule binding site for PtdInsP₂ exists on profilin.

Discussion

Our crystallographic analysis presents an atomic-resolution view of the profilin- β -actin ribbon as a discrete structural entity. Model building, constrained by diffraction data and augmented by biophysical studies of mutated actin molecules⁴³, will help to establish whether the actin-actin contacts in the ribbon are related to those in F-actin. Profilin serves as a molecular wedge between actin molecules in the ribbon, apparently binding two complementary sites on actin. The ability of profilin to bind two actin molecules suggests that profilin serves a more intricate role in regulating actin than serving simply as a monomer-sequestering protein. □

Received 11 June; accepted 24 September 1993.

- Stossel, T. B. et al. *A. Rev. Cell Biol.* **1**, 353–402 (1985).
- Pollard, T. D. & Cooper, J. A. *Rev. Biochem.* **55**, 987–1035 (1986).
- Probst, E. & Lüscher, F. *Biochim. biophys. Acta* **278**, 577–584 (1972).
- Bray, D. & Thomas, C. J. *molec. Biol.* **105**, 527–544 (1976).
- Lazarides, E. & Lindberg, U. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4742–4746 (1974).
- Lindberg, U. *J. biol. Chem.* **241**, 1246–1248 (1966).
- Carlsson, L. et al. *J. molec. Biol.* **105**, 353–366 (1976).
- Carlsson, L., Nyström, L.-E., Sundkvist, I., Markey, F. & Lindberg, U. *J. molec. Biol.* **115**, 465–483 (1977).
- Markey, F., Persson, T. & Lindberg, U. *Cell* **23**, 145–153 (1981).
- Southwick, F. S. & Young, C. L. *J. Cell Biol.* **110**, 1965–1973 (1990).
- Tilney, L. G., Bonder, E. M., Coluccio, L. M. & Mooseker, M. S. *J. Cell Biol.* **97**, 112–124 (1983).
- Safer, D., Golla, R. & Nachmias, V. T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2536–2540 (1990).
- Lassing, I. & Lindberg, U. *Nature* **314**, 472–474 (1985).
- Goldschmidt-Clermont, P. J., Machesky, L. M., Baldassare, J. J. & Pollard, T. D. *Science* **247**, 1575–1577 (1990).
- Goldschmidt-Clermont, P. J., Kim, J. W., Machesky, L. M., Rhee, S. G. & Pollard, T. D. *Science* **251**, 1231–1233 (1991).
- Lassing, I. & Lindberg, U. *J. cell. Biochem.* **37**, 255–267 (1988).
- Lassing, I. & Lindberg, U. *Exp. Cell Res.* **174**, 1–15 (1988).
- Segura, M. & Lindberg, U. *J. biol. Chem.* **259**, 3949–3954 (1984).
- Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. *Nature* **347**, 37–44 (1990).
- Schutt, C. E., Lindberg, U., Myslik, J. & Strauss, N. J. *molec. Biol.* **209**, 735–746 (1989).
- Vandekerckhove, J. S. & Weber, K. *Eur. J. Biochem.* **90**, 451–462 (1978).
- Tirion, M. T. & ben-Avraham, D. *J. molec. Biol.* **230**, 186–195 (1993).
- Bennett, W. S. & Steitz, T. A. *J. molec. Biol.* **140**, 211–223 (1980).
- Vandekerckhove, J. *Curr. Opin. Cell Biol.* **2**, 41–50 (1990).
- Holmes, K. C., Popp, D., Gebhard, W. & Kabsch, W. *Nature* **347**, 44–49 (1990).
- Noble, M. E. M., Musacchio, A., Saraste, M., Courtneidge, S. A. & Wierenga, R. K. *EMBO J.* **12**, 2617–2624 (1993).
- Janin, J. & Chothia, C. *J. biol. Chem.* **265**, 16027–16030 (1990).
- Malm, B., Larsson, H. & Lindberg, U. *J. Muscle Res. Cell Motil.* **4**, 569–588 (1983).
- Vandekerckhove, J. S., Kaiser, D. A. & Pollard, T. D. *J. Cell Biol.* **109**, 619–626 (1989).

- Carlsson, L. thesis Uppsala Univ., Sweden (1979).
- Janin, J., Miller, S. & Chothia, C. *J. molec. Biol.* **204**, 155–164 (1988).
- Rodgers, D. W., Crepeau, R. H. & Edelstein, S. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6157–6161 (1987).
- Story, R. M., Weber, I. T. & Steitz, T. A. *Nature* **355**, 318–325 (1992).
- Milligan, R. A., Whittaker, M. & Safer, D. *Nature* **348**, 217–221 (1990).
- Goldschmidt-Clermont, P. J., Machesky, L. M., Doberstein, S. K. & Pollard, T. D. *J. Cell Biol.* **113**, 1081–1089 (1991).
- Tanaka, M. & Shibata, H. *Eur. J. Biochem.* **151**, 291–297 (1985).
- Björkregren, C., Rozycki, M., Schutt, C., Lindberg, U. & Carlsson, R. *FEBS Lett.* (in the press).
- Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. *Science* **259**, 1157–1161 (1993).
- Kocks, C. et al. *Cell* **68**, 521–531 (1992).
- Theriot, J. A., Mitchison, T. J., Tilney, L. G. & Portnoy, D. A. *Nature* **357**, 257–260 (1992).
- Raghuathan, V., Mowery, P., Rozycki, M., Lindberg, U. & Schutt, C. *FEBS Lett.* **297**, 46–50 (1992).
- Yu, F.-X., Sun, H.-Q., Janmey, P. A. & Yin, H. L. *J. biol. Chem.* **267**, 14616–14621 (1992).
- Aspenström, P., Schutt, C. E., Lindberg, U. & Carlsson, R. *FEBS Lett.* **329**, 163–170 (1993).
- Winkler, F. K., Schutt, C. E. & Harrison, S. C. *Acta crystallogr.* **A35**, 901–911 (1979).
- CCP4 Package (Daresbury Laboratory, UK, 1979).
- Brünger, A. T. *X-PLOR v. 3.0 manual* (Yale Univ., New Haven, 1992).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
- Finzel, B. C. et al. in *Crystallographic Modelling Methods in Molecular Design* (eds Bugg, C. E. & Ealick, S. E.) 175–188 (Springer, New York, 1990).
- Ampe, C., Markey, F., Lindberg, U. & Vandekerckhove, J. *FEBS Lett.* **228**, 17–21 (1988).
- Myslik, J. C. thesis, Princeton Univ., USA (1992).
- Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).

ACKNOWLEDGEMENTS. We thank W. Kabsch and K. Holmes for providing us with the atomic coordinates of α -actin before their deposition in the Brookhaven Protein Database and T. Twomey and E. Girit for technical assistance. This work was supported by grants to C.E.S. from the National Institutes of Health, Carter-Wallace Inc. and Sterling-Winthrop Pharmaceuticals, and to U.L. from the Swedish Cancer Foundation, the Swedish Natural Science Research Council and the Granholm Foundation. M.D.R. was supported by a Muscular Dystrophy Association Postdoctoral Fellowship and J.C.M. was supported by an NIH Predoctoral training grant in biophysics.

A second companion of the millisecond pulsar 1620–26

D. C. Backer, R. S. Foster† & S. Sallmen*

* Astronomy Department, University of California, Berkeley, California 94720, USA

† Remote Sensing Division, Naval Research Laboratory, Code 7210, Washington 20375, USA

MILLISECOND pulsars are usually found in binary systems. This is in keeping with the generally accepted model for the formation of such pulsars¹⁴ in which an old neutron star is spun up to high angular velocities by the accretion of matter from a companion star. The millisecond pulsar 1620–26 in the globular cluster M4 is no exception: timing measurements^{1,2} reveal the presence of a 0.3-solar-mass companion star (probably a white dwarf) with an orbital period of 191 days. But subsequent measurements of this pulsar have identified a small but significant deviation from the expected behaviour^{3,4}, suggestive of an unusually large second derivative in the pulsar's rotation rate⁵. Here we examine several possible sources—both intrinsic and extrinsic—for this derivative, and we find that it is best explained by the presence of a second, weakly bound companion object moving in a wide orbit around the main binary system. The third body in this hierarchical system has an orbital period of ~100 years and a mass approximately ten times that of Jupiter, and may have been captured during a recent collision with another stellar system⁶.

Since January 1988 we have conducted signal arrival time observations of PSR 1620–26 at 800 MHz and 1330 MHz, at approximately two-month intervals, with the NRAO (National Radio Astronomy Observatory) 42-m telescope. Details of these measurements are presented elsewhere^{7,8}. The data were analysed using a version of the standard timing data analysis program TEMPO⁹. We transformed the data to the frame of the pulsar by removing a model of the Earth's motion about the Solar

System barycentre and by adding a model of the pulsar's motion about the binary system barycentre. In the frame of the pulsar the data can be modelled by a polynomial function of time with terms for phase, spin frequency Ω , and spin-down rate $\dot{\Omega}$. The residuals from a fit of the PSR 1620–26 data to the ten parameters of this model are dominated by a large cubic residual that has a peak amplitude of one spin period (Fig. 1a). A fit for a second derivative of the spin frequency, $\ddot{\Omega}$, removes this cubic phase residual (Fig. 1b). The second derivative of the fit is $\ddot{\Omega} = +1.2 \times 10^{-22} \text{ s}^{-3}$. An independent fit to our data which included a proper motion term is consistent with the Cudworth and Hanson¹⁰ value, but has low statistical significance at present. The parameters of this fit are given in Table 1. Consistent parameters from separate observations have been obtained by A. Lyne (personal communication) and by Thorsett *et al.*¹¹. The low amplitude systematic deviations evident in Fig. 1b can be removed by a simultaneous fit for a third derivative of the spin rate. Because this additional parameter is strongly coupled with other spin parameters, we will use three times the fit value, $\ddot{\Omega} = 1.3 \times 10^{-31} \text{ s}^{-4}$, as an upper limit. We discuss next the likelihood of several sources of the large second derivative reported for this pulsar.

Intrinsic origins. Pulsar spin and magnetic axes are not aligned, inducing a torque that slowly brakes all isolated stars. The assumption of pure magnetic dipole radiation as the source of the torque leads to a unique value for the change in the electromagnetic spin-down rate, $\ddot{\Omega}_{\text{EM}}$. In the case of constant magnetic moment and constant moment of inertia (for example, ref. 12),

$$\ddot{\Omega}_{\text{EM}} = \frac{3\dot{\Omega}_{\text{EM}}^2}{\Omega_{\text{EM}}} = +1.0 \times 10^{-29} \text{ s}^{-3}.$$

The Ω quantities are subscripted with EM to distinguish them from the measured quantities, which may be affected by dynamical effects. The observed $\ddot{\Omega}$ is larger than $\ddot{\Omega}_{\text{EM}}$ by seven orders of magnitude.

The observed $\ddot{\Omega}$ could result from other effects internal to the neutron star, such as timing noise, but the lack of any similar effect in other millisecond pulsars is a strong argument against this explanation. In slowly rotating pulsars, occasionally $\ddot{\Omega} \gg \ddot{\Omega}_{\text{EM}}$, but these systems have a low frequency spectrum of timing noise that is not evident in PSR 1620–26.

Dynamical origins. The large second derivative of the rotation rate of PSR 1620–26 could be the result of an unmodelled motion of the inertial reference frame of the pulsar–companion system. Low order dynamical effects—position, velocity and

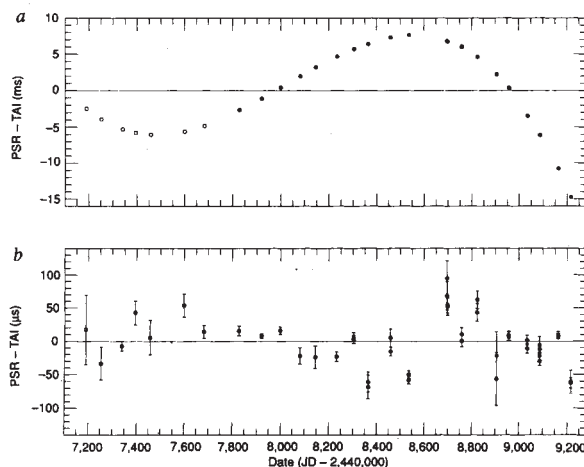


FIG. 1 Residuals between the pulse arrival times (PSR) observed at 1.3 GHz and models calculated in international atomic time (TAI). Open and filled symbols distinguish the two data sets, which were obtained with independent hardware systems; an arbitrary offset between the two sets has been removed. The errors plotted are estimated from internal dispersion of measurements. Multiple data samples at later epochs are the result of independent data from signals spaced by 20 MHz. A time-variable interstellar dispersion is removed using measurements at 0.8 GHz (ref. 8). In a the model does not include a second derivative of the rotation rate and the spin parameters are adjusted to minimize r.m.s. residuals. Errors are negligible on this scale. In b the model includes the second derivative with the best fit and an assumed proper motion which has been determined (for the cluster centroid) by optical astrometry¹⁰.

TABLE 1 Parameters of PSR 1620–26

Parameter	Value	Uncertainty
Right ascension (B1950.0)	16 h 20 min 34.1607 s	0.0007 s
Declination (B1950.0)	−26° 24′ 58.08″	0.05″
Period, $2\pi/\Omega$	0.0110757508045800 s	0.0000000000000140 s
Period derivative	$0.97744 \times 10^{-18} \text{ s s}^{-1}$	$0.00040 \times 10^{-18} \text{ s s}^{-1}$
Period second derivative	$-2.3015 \times 10^{-27} \text{ s s}^{-2}$	$0.0040 \times 10^{-27} \text{ s s}^{-2}$
Epoch	2447188.00415 JD	
Dispersion measure	62.303 pc cm ^{−3}	
Projected semi-major axis $a_1 \sin i$	64.8094959 s	0.0000100 s
Eccentricity, e	0.02531612	0.00000040
Epoch of periastron, T_0	2447197.22094 JD	0.00040 JD
Binary period, P_b	16540660.213 s	0.14 s
Longitude of periastron, ω	117.13119°	0.00070°

Uncertainties quoted are twice the formal error in the fit. The Center for Astrophysics ephemeris 740R is used. Variations of the dispersion measure with amplitude of 0.005 pc cm^{-3} are detected and removed from individual epochs of observations. The fit above assumes the absolute proper motion given by Cudworth and Hanson¹⁰: (−11.6, −15.7) milliarcsec yr^{−1}. A constant offset is removed between data sets with different hardware at JD 2447700 with an estimated accuracy of 25 μs.

acceleration—cannot be distinguished from small shifts in the intrinsic spin parameters of the pulsar. The measured second derivative of the spin frequency allows us to calculate an apparent first derivative of the acceleration (a second derivative of the rotation frequency) of the pulsar–companion system. We refer to this as a jerk, $\dot{a} = c\dot{\Omega}/\Omega = +6.2 \times 10^{-5} \text{ cm s}^{-3}$, where c is the speed of light. If the measured second derivative of the rotation frequency persists, then the first derivative will change sign in ten years and will indicate the dominance of the effect of acceleration over that of the intrinsic spin-down rate. Negative values of $\dot{\Omega}$, indicative of similar dynamical effects, are found for two pulsars in the dense core of the M15 globular cluster¹³.

In the context of models where millisecond pulsars are spun up by angular momentum transfer from a companion star (for example, ref. 14), we can estimate a range of accelerations of the system. In these models, $\dot{\Omega}_{\text{EM}}$ could be as large as $-1.3 \times 10^{-13} \text{ s}^{-2}$, if this pulsar was spun-up recently to the maximum value, or as small as $-9.7 \times 10^{-16} \text{ s}^{-2}$ if it is as old as the cluster. The corresponding extremes for accelerations are $+4.0 \times 10^{-6}$ and $-2.6 \times 10^{-6} \text{ cm s}^{-2}$.

The inferred jerk could be the result of the pulsar's orbit in the smoothed central gravitational potential of M4, which can be estimated from the known cluster parameters and the pulsar's position. The typical acceleration is $a_{\text{gc}} = 4\pi G\rho_c b/3 = 10^{-6} \text{ cm s}^{-2}$, where the core density ρ_c is $4.1 \times 10^4 M_{\odot} \text{ pc}^{-3}$ (ref. 15) and $b = 44''$ (0.4 pc) is the distance of the pulsar system relative to the cluster centre. This acceleration is within the bounds discussed above. The typical jerk is $\dot{a}_{\text{gc}} = 4\pi G\rho_c \sigma_c/3 = 3.1 \times 10^{-19} \text{ cm s}^{-3}$, where the velocity dispersion σ_c is 4.2 km s^{-1} (ref. 15). This jerk is four orders of magnitude too small. Although this conclusion is based on circular orbits, consideration of radial orbits does not result in a significant additional factor (for example, ref. 16).

Phinney¹⁶ has explored the possibility that the inferred jerk is the result of an encounter with a nearby $M < 0.5 M_{\odot}$ (solar mass) star. The probability that such a star passes close enough to produce the measured $\dot{\Omega}$ is only $\sim 2 \times 10^{-5} (M_{\odot}/M)^{1/2}$. In such an encounter the contribution to $\dot{\Omega}$ is minimized if the passing star's velocity and jerk are nearly parallel to the line of sight. The observed limit on the acceleration would require alignment to within a few degrees. Nearest neighbour dynamical effects are not a likely source of the observed effect.

We next consider an explanation in the form of a second bound companion with an orbit much larger than that of the pulsar's first companion, making this a hierarchical triple system. The inferred jerk and the range of possible accelerations presented above allow us to explore the range of possible masses and periods of this third body. Although the low binding energy of this third companion would make an elliptical orbit most probable, we first consider only circular orbits. Figure 2 displays the minimum mass and period, derived from Kepler's law for a circular orbit of inclination 90° , which yields the inferred jerk. The large and positive jerk ($\dot{\Omega}$) and small acceleration ($\dot{\Omega}$) indicate that the third body in this model is about to cross the line of nodes (with velocity away from the observer). This unique phase in the orbit favours shorter orbital periods in the calculation.

We can constrain the mass and orbital period of the second companion further by requiring that the first derivative of the acceleration be consistent with the upper limit on $\dot{\Omega}$ discussed above. We investigated a set of circular orbits with inclination angle 45° and a range of masses and periods that are consistent with all the observed effects during five-year segments of the trial orbits. Figure 2 displays the successful masses and periods above the minimum values of 80 yr and 10^{31} g which come from the upper limit on $\dot{\Omega}$. This mass is five times that of Jupiter. The ratio of the first and third derivatives of Ω would provide an exact measure of the orbital period for a circular orbit. Unfortunately, we have only upper limits for both quantities. The ratio of the limits yields an estimated period of 120 yr. The

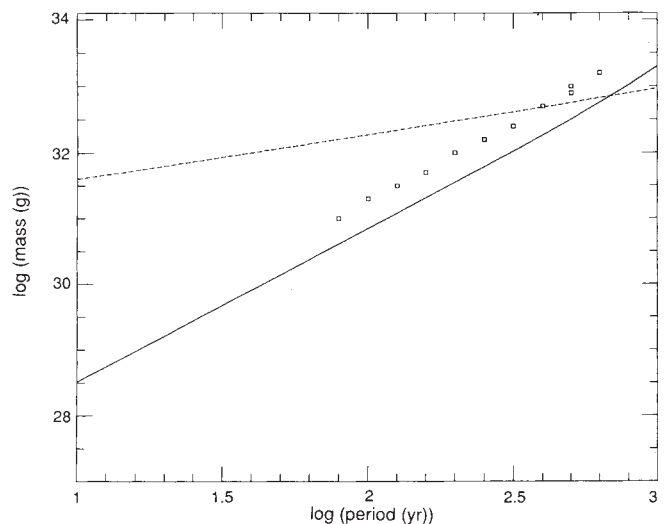


FIG. 2 Estimates of companion mass and orbital period which could explain by a dynamical effect the large-period second derivative reported for PSR1620–26 (an acceleration first derivative or jerk). The solid line is the minimum mass for a given period that would produce the inferred jerk. Open symbols are values of mass and period that produce the inferred jerk and are consistent with the limits on acceleration and first derivative of the acceleration discussed in the text. Circular orbits with an inclination of 45° are assumed. The dashed line divides higher-mass second companions whose binding energies exceed the typical kinetic energy of stars in the cluster from lower-mass companions that are weakly bound.

corresponding mass is $2 \times 10^{31} \text{ g}$. After the outer mass has moved through a significant fraction of its orbit we will determine the keplerian parameters of the two orbits as well as any tidal influence of the outer body on the inner orbit. The lower-mass models which have been investigated in detail by Sigurdsson⁶ require favourable values of eccentricity and longitude of periastron to keep the expected value of $\dot{\Omega}$ below our current limit.

The survival of this outer companion in such a wide orbit in the core of M4 is problematic because of the low binding energy and frequent encounters with cluster stars. Plotted in Fig. 2 is a line that separates the close binaries that survive cluster crossings and in fact become more closely bound in collisions—from binaries where the companions are more separated, and thus have a high probability for disruption. The collision time for the disruption of the outer orbit, during which this pulsar system crosses the cluster hundreds of times, is of the order of 10^8 yr , which is also the lower limit to the age of the pulsar. The most tightly bound (that is, highest mass, longest orbital period) companion that is consistent with earlier limits then best explains the observed second derivative. Alternatively, lower masses^{6,15} are possible if the third body was acquired in a relatively recent encounter with another stellar system. We can look forward to improved estimates of the parameters of the second orbit and further investigations for formation models. □

Received 18 May; accepted 24 August 1993.

1. Lyne, A. G. et al. *Nature* **332**, 45–49 (1988).
2. McKenna, J. & Lyne, A. G. *Nature* **336**, 226–227 (1988); erratum **336**, 698 (1988).
3. Foster, R. S. thesis, Univ. of California at Berkeley (1990).
4. Thorsett, S. E. thesis, Princeton Univ. (1991).
5. Backer, D. C. in *Planets Around Pulsars*, *Astr. Soc. Pacif. Conf. Ser.* 36 (eds Phillips, J. A., Thorsett, S. E. & Kulkarni, S. R.) 11–18 (Astr. Soc. Pacif., San Francisco, 1993).
6. Sigurdsson, S. *Astrophys. J.* (in the press).
7. Foster, R. S., Fairhead, L. & Backer, D. C. *Astrophys. J.* **378**, 687–695 (1991).
8. Backer, D. C., Hama, S., Van Hook, R. S. & Foster, R. S. *Astrophys. J.* **404**, 636–642 (1993).
9. Taylor, J. H. & Weisberg, J. M. *Astrophys. J.* **345**, 434–450 (1989).
10. Cudworth, K. M. & Hanson, R. D. *Astr. J.* **105**, 168–172 (1993).
11. Thorsett, S. E., Arzoumanian, Z. & Taylor, J. H. *Astrophys. J.* **412**, L33–L36 (1993).
12. Michel, F. C. *Theory of Neutron Star Magnetospheres* (Univ. of Chicago Press, 1991).
13. Anderson, S. B. thesis, California Inst. of Technology (1993).

14. van der Heuvel, E. P. J. in *Planets Around Pulsars*, Astr. Soc. Pacif. Conf. Ser. 36 (eds Phillips, J. A., Thorsett, S. E. & Kulkarni, S. R. 123–147 (Astr. Soc. Pacif., San Francisco, 1993).
15. Trager, S. C., Djorgovski, S. G. & King, I. R. in *Structure and Dynamics of Globular Clusters*, Astr. Soc. Pacif. Conf. Ser. 50 (eds Djorgovski, S. G. & Meylan, G.) 347–355 (Astr. Soc. Pacif., San Francisco, 1993).
16. Phinney, E. S. in *Planets Around Pulsars*, Astr. Soc. Pacif. Conf. Ser. 36 (eds Phillips, J. A., Thorsett, S. E. & Kulkarni, S. R.) 141–169 (Astr. Soc. Pacif., San Francisco, 1993).

ACKNOWLEDGEMENTS. We thank A. Fruchter, A. Lyne, S. Sigurdsson and S. Thorsett for discussions. The 42-m telescope is operated by the National Radio Astronomy Observatory under a cooperative agreement with the US National Science Foundation.

The origin of Pluto's peculiar orbit

Renu Malhotra

Lunar and Planetary Institute, 3600 Bay Area Boulevard,
Houston, Texas 77058, USA

THE origin of Pluto's unusual orbit—the most eccentric and inclined of all the planets—remains a mystery. The orbits of Pluto and Neptune overlap, but close approaches of these two planets are prevented by the existence of a resonance condition¹: Pluto's orbital period is exactly 3/2 that of Neptune, which ensures that the conjunctions always occur near Pluto's aphelion. Long-term orbit integrations^{2–5} have uncovered other subtle resonances and near-resonances, and indicate that Pluto's orbit is chaotic yet remains macroscopically stable over billion-year timescales. A suggestion⁴ that the orbit may have evolved purely by chaotic dynamics appears unlikely in light of recent orbital stability studies⁶, unless one appeals to a well-timed collision to place Pluto in its stable orbit⁹. Here I show that Pluto could have acquired its current orbit during the late stages of planetary accretion, when the jovian planets underwent significant orbital migration as a result of encounters with residual planetesimals⁷. As Neptune moved outwards, a small body like Pluto in an initially circular orbit could have been captured into the 3:2 resonance, following which its orbital eccentricity would rise rapidly to its current Neptune-crossing value.

Consider a mass m_c of planetesimals ejected to Solar System escape orbits by a planet of mass M at orbital radius r . From conservation of angular momentum, it follows that the planet suffers a decrease of orbital angular momentum and a concomitant decrease of orbital radius given by

$$\delta r/r \approx -2(\sqrt{2}-1)m_c/M \quad (1)$$

For planetesimals that are scattered outwards away from the Solar System but do not achieve escape orbits (remaining bound in the Oort cloud of comets which exists as a halo surrounding our planetary system at distances greater than 20,000 astronomical units (AU)), the numerical factor in equation (1) would be slightly smaller. Conversely, planetesimals scattered into the inner Solar System increase the orbital radius and angular momentum of the planet. The four jovian planets acting together, however, behave counter to this simple picture. Consider planetesimal scattering by Neptune: those planetesimals scattered outwards end up in the Oort cloud, or return to be re-scattered. A fraction of the latter set is scattered inwards. The inwardly scattered planetesimals are systematically ejected from the planetary system by the more massive Jupiter. Thus the outer, less massive planets, Neptune and Uranus, transfer control of the ejection process to the inner, more massive Jupiter and Saturn. (Note that Pluto's small mass—only two-thousandths of the Earth's mass—makes it an ineffective scatterer of planetesimals.) Numerical simulations of this process show that Neptune, Uranus and Saturn suffer a net gain of angular momentum while Jupiter suffers a net loss⁷.

The timescale and extent of the radial migration of the planets is determined by the details of the late stages of the planet formation process and is poorly constrained by current theories. The mass of the Oort cloud, which is believed to have been populated by small bodies ejected from the zones of the jovian planets, is estimated to be $\sim 10 M_\oplus$ ($M_\oplus$ = mass of the Earth) (ref. 8). The ejection of such a mass of planetesimals corresponds to an inward migration of Jupiter of a few tenths of an AU, and an outward migration of Neptune of several AU. In the final stages of planet formation, ejection and accretion of planetesimals will occur concurrently; therefore, the timescale for the radial migration of the planets is expected to be comparable to that of accretion. An approximate lower limit for the latter is the observed $\sim 10^7$ yr lifetime of dust in disks around pre-main-sequence stars⁹.

One consequence of Neptune's orbital expansion is that its orbital resonances would have swept across a range of heliocentric distances comparable to its radial migration. During this resonance sweeping, a small body such as Pluto, initially in a nearly circular orbit beyond Neptune, would have a significant probability of being captured into a resonance. Resonant perturbations from Neptune would transfer sufficient angular momentum to the body to keep it trapped in the resonance and to expand its orbit in concert with Neptune. A by-product of such a resonance capture is that Pluto's orbital eccentricity would have been excited to a high value. This may be seen by considering the perturbation equations in the restricted three-body approximation¹⁰. Near a $j:j+1$ resonance (where j is an integer), the resonant perturbations from Neptune on Pluto's orbital frequency, n_P , and eccentricity, e_P , are described by the following equations:

$$\begin{aligned} \dot{n}_P &= 3(j+1)\mu_N n_P^2 e_P f(\alpha) \sin \phi \\ \dot{e}_P &= -\mu_N n_P f(\alpha) \sin \phi \end{aligned} \quad (2)$$

where the 'resonance angle' is defined by

$$\phi = (j+1)\lambda_P - j\lambda_N - \bar{\omega}_P \quad (3)$$

λ denotes the mean longitude, $\bar{\omega}$ is the longitude of perihelion, μ_N is the mass of Neptune relative to the Sun, $\alpha = a_N/a_P < 1$ is the ratio of the semimajor axes of Neptune and Pluto, and $f(\alpha)$ is a positive function, with numerical value 2.48 at the 3:2 resonance. Let $\langle \dot{n}_N \rangle$ be the mean rate of change of Neptune's orbital frequency as its orbit expands. Upon capture into resonance, the orbital frequency of Pluto becomes locked to that of Neptune, that is

$$(j+1)\langle \dot{n}_P \rangle \approx j\langle \dot{n}_N \rangle \quad (4)$$

It then follows from equations (2) and (4) that

$$\left\langle \frac{de_P^2}{dt} \right\rangle \approx -\frac{2}{3(j+1)} \left\langle \frac{\dot{n}_N}{n_N} \right\rangle = \frac{1}{(j+1)} \left\langle \frac{\dot{a}_N}{a_N} \right\rangle \quad (5)$$

where the last expression follows from the keplerian relation between the orbital frequency and the semimajor axis ($n^2 a^3 = \text{constant}$). In other words, upon capture into resonance, Pluto's eccentricity would increase at a rate determined by the average rate of expansion of Neptune's orbit. Equation (5) can be integrated to yield

$$e_{P,\text{final}}^2 - e_{P,\text{initial}}^2 \approx \frac{1}{j+1} \ln \left(\frac{a_{N,\text{final}}}{a_{N,\text{initial}}} \right) \quad (6)$$

Neptune's current semimajor axis is $a_{N,\text{final}} = 30.17$ AU. Thus, we deduce that in order to excite Pluto's eccentricity from near-zero to 0.25, $a_{N,\text{initial}} \approx 25$ AU is required at the time of Pluto's capture into the 3:2 Neptune resonance ($j=2$ in equation (6)); Pluto's initial orbital radius would have been 32.8 AU. Remarkably, this result is independent of the timescale (and other details) of the expansion of Neptune's orbit. But there exists a practical constraint on the orbit expansion timescale, namely that it be much

longer than the period of the resonant perturbations, in order that resonance capture be effected¹¹. This lower limit on the orbit expansion timescale is $\sim 10^5$ yr.

The above analysis is based on first-order perturbation theory, and it is legitimate to question its validity, particularly in view of the fact that at the high eccentricity of 0.25, Pluto's orbit would become Neptune-crossing. Furthermore, the perturbations from the other jovian planets on Pluto (as well their mutual interactions) have been neglected in this analysis. In order to test the hypothesis more rigorously, we have carried out a numerical integration of the orbit evolution of the four jovian planets and Pluto. As Pluto's mass is several orders of magnitude smaller than the jovian planets, we treated it as a massless 'test particle'. The numerical method used was a modified version of the mixed-variable-symplectic method¹²⁻¹⁴, which is a very fast integrator particularly suited to Solar System integrations.

I assumed a simple model for the time variation of the orbital semimajor axes of the outer planets:

$$a(t) = a_f - \Delta a e^{-t/\tau} \quad (7)$$

where a_f is the value at the current epoch and τ is the orbit expansion timescale. I chose $\Delta a = 0, 1, 3$ and 6 AU for Jupiter, Saturn, Uranus and Neptune, respectively, consistent with the net migration induced by the ejection of a few $M_\oplus$ of planetesimals from the outer planetary region. This orbital migration was modelled in the equations of motion by means of a force of magnitude $f = f_0 e^{-t/\tau}$ on each planet along the direction of its velocity. (f_0 can be related in a straightforward manner to a_f , Δa and τ .) So that the simulation would use a reasonable amount of computer time, the total integration time was 20 Myr, and τ was chosen to be 1.5 Myr. The qualitative behaviour is expected to be insensitive to τ for $\tau \geq 10^5$ yr, which ensures an approximate adiabatic condition for the evolution near the 3:2 Neptune resonance.) Twenty 'test particle' Plutos were integrated along with the planets. Their initial orbital elements were chosen randomly as follows: a in the range 32.5–33.5 AU (just outside the 3:2 Neptune resonance), e in the range 0.0–0.3, i in the range 0° – 10° , and the angles (mean longitude, longitude of perihelion and ascending node) in the range 0 – 2π .

A typical case of the evolution found in the simulations is shown in Fig. 1. Nineteen of the 20 test Plutos were captured in the 3:2 resonance and the evolution proceeded in a manner very similar to that predicted by the simple analysis. (The one test Pluto that did not follow this evolution was captured in a 4:3 resonance and eventually had a close encounter with Neptune.) The behaviour of Pluto's a and e varied very little amongst the 19 cases, but the inclination evolution was rather sensitive to the initial conditions. The behaviour of the resonance angle ϕ was very stable; the final amplitude of its oscillations was found to be distributed in the range 30° – 100° . (The observed value is 76° ; recent numerical integrations¹⁵ show that the long-term stability of the resonance requires a value $< 80^\circ$.) We conclude that the analytical argument above does capture the salient features of the dynamical evolution. The long-period variations, evident particularly in the inclination, are possibly due to a nearby inclination-type secular resonance¹⁶. Longer integrations are required to determine the range of behaviours possible for Pluto's inclination.

A few words about the evolution of the jovian planets are in order. Although the initial orbits of the planets were significantly different from their current orbits, the behaviour of the orbital eccentricities and inclinations found in the simulations was similar to that observed in the long time integrations of the current planetary configuration³.

The scheme outlined here for the origin of Pluto's highly eccentric, Neptune-crossing orbit is quite robust in that the probability of capturing Pluto in the 3:2 Neptune resonance from an initially circular orbit is very high. However, there are a few

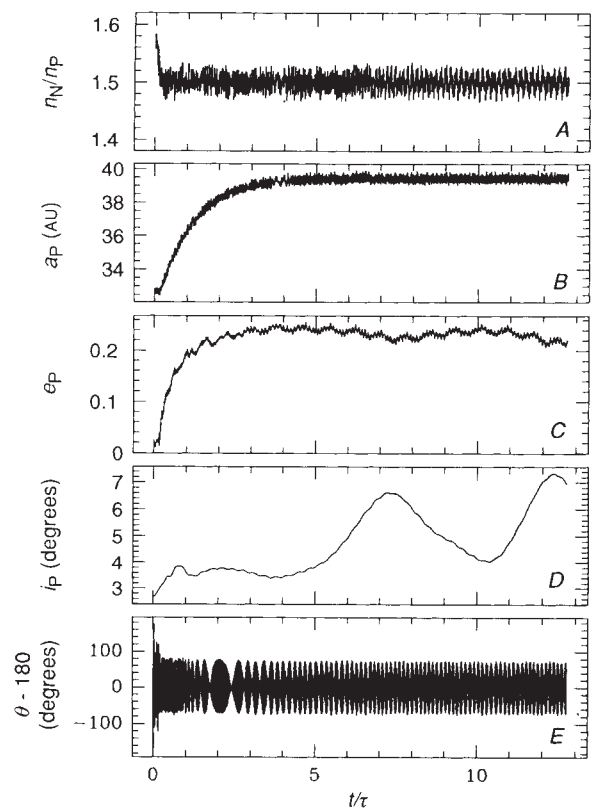


FIG. 1 The time evolution of Pluto's orbital elements is shown for a typical case of the capture of Pluto into the 3:2 Neptune resonance. (The time is indicated in units of τ , the characteristic timescale for the expansion of Neptune's orbit.) As Neptune's orbit evolves outward, the mean motion ratio (A) decreases initially until it reaches the value 3/2; at this point Pluto is captured in the resonance and subsequently its orbit also expands (B) in concert with Neptune. Pluto's eccentricity is pumped up rapidly (C) to a Neptune-crossing value. The evolution of the inclination (D) appears to be dominated by a nearby secular resonance, but remains to be understood. The resonance angle goes from circulations to remarkably constant-amplitude librations (E).

points of note. (1) I have modelled the radial migration of the jovian planets as a smooth process, whereas in reality it is stochastic. We expect that, as long as no single event produces too large a change in the planetary orbits, the smooth approximation should work well. A rough estimate for the mass of an ejected planetesimal that would invalidate this assumption can be obtained by equating the change in the orbital radius, $|\delta r/r|$ (Equation 1), with the half-width of the 3:2 Neptune resonance, $\delta a_P/a_P \approx 4(\mu_N f(a) e_P/3)^{1/2}$ (ref. 11); this yields a mass $m_c \sim 10^{27}$ g, which is two orders of magnitude greater than the mass of Pluto. Thus, if we consider that Pluto was among the largest members of the planetesimal population beyond the orbit of Neptune surviving to the evolutionary stage considered here, then the smooth approximation appears quite reasonable. (2) In the numerical simulations, I assumed that the masses of the jovian planets were fixed at their current values during the entire evolution. The probability of resonance capture will depend upon the mass of Neptune. However, for sufficiently slow expansion of the planet orbits, resonance capture theory shows that the dependence on the mass of the perturbing planet is not very strong¹¹. Therefore this limitation of the simulations is not likely to be of great import. (3) The analysis given here shows that the most direct requirement is on Neptune's orbital migration: $\Delta a_N \geq 5$ AU. The sensitivity of the dynamics to the Δa s of the other planets remains to be determined. The values chosen in

my simulation ensure that during their radial migration, the jovian planets do not encounter any strong orbital resonances amongst themselves, and therefore suffer only relatively small mutual perturbations. How restrictive this condition might be on the entire dynamical evolution described here requires further study. (4) Finally, the role of possible planetesimal collisions with Pluto during its evolution in the 3:2 Neptune resonance also needs to be evaluated. This may have an important bearing on the origin and properties of the Pluto-Charon binary.

The significant orbital evolution of the jovian planets implied by the model outlined here would have implications for a number of other problems concerning the outer Solar System, such as the capture of Triton by Neptune¹⁷ and the structure of the putative Kuiper belt¹⁸, as well as for studies of the evolution of the asteroid belt. □

Received 5 July; accepted 6 September 1993.

1. Cohen, C. J. & Hubbard, E. C. *Astr. J.* **70**, 10–13 (1965).
2. Williams, J. G. & Benson, G. S. *Astr. J.* **76**, 167–177 (1971).

3. Applegate, J. H., Douglas, M. R., Gursel, J., Sussman, G. J. & Wisdom, J. *Astr. J.* **92**, 176–194 (1986).
4. Sussman, G. J. & Wisdom, J. *Science* **241**, 433–437 (1988).
5. Milani, A., Nobili, A. M. & Carpino, M. *Icarus* **82**, 200–217 (1989).
6. Levison, H. F. & Duncan, M. J. *Astrophys. J.* **406**, L35–L38 (1993).
7. Fernandez, J. A. & Ip, W. H. *Icarus* **58**, 109–120 (1984).
8. Weissman, P. R. *Nature* **344**, 825–830 (1990).
9. Strom, S. E., Edwards, S. & Strom, K. M. in *The Formation and Evolution of Planetary Systems* (eds Weaver, H. A. & Danly, L.) 91–107 (Cambridge Univ. Press, 1989).
10. Brouwer, D. & Clemence, G. M. *Methods of Celestial Mechanics* (Academic, New York, 1961).
11. Dermott, S. F., Malhotra, R. & Murray, C. D. *Icarus* **76**, 295–334 (1988).
12. Wisdom, J. & Holman, M. *Astr. J.* **102**, 1528–1538 (1991).
13. Saha, P. & Tremaine, S. *Astr. J.* **104**, 1633–1640 (1992).
14. Malhotra, R. *Celestial Mech. and Dynamical Astr.* (submitted).
15. Levison, H. F. & Stern, S. A. *Lunar planet. Sci.* **24**, 869–870 (1993).
16. Knezevic, Z., Milani, A., Farinella, P., Froeschle, Ch. & Froeschle, Cl. *Icarus* **93**, 316–330 (1991).
17. McKinnon, W. B. *Nature* **311**, 355–358 (1984).
18. Duncan, M. J. & Quinn, T. A. *Rev. Astr. Astrophys.* **31**, 265–295 (1993).
19. Levison, H. F. & Stern, S. A. in *Pluto-Charon Conf.*, 6–9 July 1993, Flagstaff, Arizona, USA.

ACKNOWLEDGEMENTS. I thank S. Dermott, A. Milani and A. Nobili for discussion, M. Duncan for computer programs, and P. Goldreich for comments on a preliminary draft of this letter. This work was supported by NASA.

Conducting charge-transfer salts based on neutral π -radicals

C. D. Bryan*, A. W. Cordes*, R. M. Fleming†, N. A. George‡, S. H. Glarum†, R. C. Haddon†, R. T. Oakley†, T. T. M. Palstra†, A. S. Perel†, L. F. Schneemeyer† & J. V. Waszczak†

* Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701, USA

† AT&T Bell Laboratories, Murray Hill, New Jersey 07974-0636, USA

‡ Guelph-Waterloo Centre for Graduate Work in Chemistry, Guelph Campus, Department of Chemistry and Biochemistry, University of Guelph, Guelph, Ontario N1G 2W1, Canada

MOST molecular conductors rely on charge transfer to create carriers. For example, the ET salts¹ are hole-doped whereas the C_{60} salts² are electron-doped. Neutral radical species in which bands are formed by π -orbital overlap would be expected to have half-filled bands and thus to be conducting³, but no such metals have yet been reported. Here we report the synthesis and characterization of a molecular conductor which combines both of these approaches: energy bands are formed from one-dimensional stacks of neutral π -radicals, and the material is rendered conducting by electron transfer from the conduction band following doping with an acceptor. The radical species is the 1,4-phenylene-bis(dithiadiazolyl) diradical 1,4-[(S₂N₂C)₆H₄(CN₂S₂)] (2 in Fig. 1), reaction of which with iodine vapour leads to crystals of [2][I]. At low temperatures this compound is essentially a diamagnetic insulator, but above 200 K the conductivity and magnetic susceptibility increase markedly, and at room temperature the conductivity reaches 100 S cm⁻¹, which is comparable to that shown by conventional molecular charge-transfer salts.

The preparation of synthetic conductors usually involves the oxidation or reduction of a closed-shell molecule to generate a mixed-valence or charge-transfer salt^{1,2}. We have been interested in an alternative approach, namely the use of neutral π -radicals, in which the requirement for a partially filled energy band is inherently fulfilled³. There are, however, few π -radicals with sufficient stability and flexibility^{4–6}. Initially we examined variants of the phenalenyl radical^{7–10}, but recent advances in heterocyclic chemistry have provided a variety of potentially useful π -radical building blocks^{11,12}. One objective has been to generate

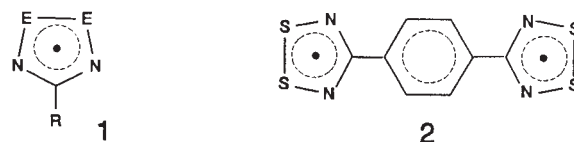


FIG. 1 The molecular structures of the neutral radical dithiadiazolyl (1), where E is S or Se, and the 1,4-phenylene-bis(dithiadiazolyl) diradical (2).

materials consisting of molecular stacks or columns, and for several choices of the R-group in 1 (Fig. 1, E = S, Se), this packing arrangement has been realized^{11,13}. By definition, however, neutral radicals give rise to a half-filled electronic energy band, which in a one-dimensional packing arrangement is prone to a Peierls instability (essentially a dimerization process) driven by charge density waves (CDWs). Because the radicals tend to dimerize, we have attempted to stabilize the metallic state by applying pressure and by preparing materials in which two- and three-dimensional interactions are enhanced. Several small-band-gap semiconductors of this kind (with E = Se) have been characterized^{11,13,15}, but no truly metallic compound has yet been realized.

An alternative strategy for stabilizing the metallic state involves doping the energy band away from the half-filling which is characteristic of the neutral state. We have now found that the oxidation of the 1,4-phenylene-bridged bifunctional radical 2 (Fig. 1, ref. 14) with iodine produces a highly conductive mixed-valence salt of formula [2][I]. The preparation requires heating a mixture of 2 (1.0 mmol) and I₂ (0.5 mmol) in a sealed, evacuated (1 mTorr) pyrex tube (25 × 250 mm). Slow fractional sublimation along a temperature gradient (225 °C to 160 °C over 100 mm) affords, after several days, lustrous silver black needles with elemental composition C₈H₄N₄S₄I (yield ~70%).

The crystal structure (Fig. 2) consists of columns of uniformly spaced (3.415(2) Å) molecules of 2 interspersed by columns of disordered iodines running along the x-axis. Perpendicular to the stacking direction the heterocyclic rings lie in dove-tailed arrays running parallel to the y-axis. Within the CN₂S₂ rings, the S–S (2.066(1) Å) and S–N (1.616(2) Å) bonds are between those seen in the neutral dimer of 2 (ref. 14) and its di-cation [2]²⁺ (ref. 15), but closer to the former, implying an oxidation state (per ring) of <0.5.

The magnetic susceptibility of [2][I] was determined between 4 K and the decomposition point using the Faraday technique (Fig. 3). At low temperatures there is a small Curie tail due to

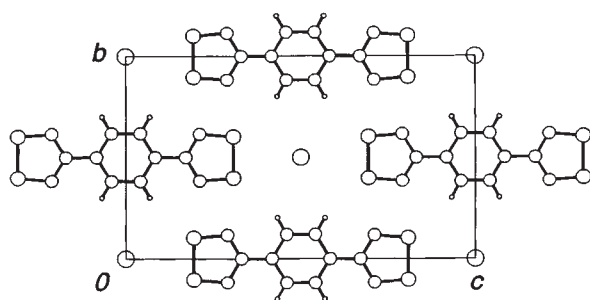


FIG. 2 The crystal structure of $[1,4-(S_2N_2C)C_6H_4(CN_2S_2)][I]$. The space group is orthorhombic $Immm$ with (at 23 °C) $a=3.415(2)$ Å, $b=10.061(2)$ Å, $c=16.740(2)$ Å, $FW=411.31$, $Z=2$. (Where FW is the formula weight, and Z is the number of molecules in the unit cell.) The crystals have a large mosaic spread, typically $\sim 1^\circ$. The non-hydrogen coordinates are S, 0, 0.10269(5), 0.19337(3); N, 0, 0.11573(20), 0.28955(12); C1, 0, 0, 0.32886(19); C2, 0, 0, 0.41657(18); C3, 0, 0.11905(23), 0.45886(14). The iodine atoms are disordered along x . In the final room temperature refinement, (R -value, 0.027) the iodine positions were modelled by partial iodines at 0, 0, 0 (occupancy 0.472), ± 0.225 , 0, 0 (0.221) and 0.500, 0, 0 (0.104).

defects in the crystal. Above 180 K there is a steep rise in the paramagnetism, which becomes almost linear above room temperature. An electron spin resonance study of crystals of $[2][I]$ showed the same rise in susceptibility, with a large increase in the linewidth from 20 G at 100 K to 130 G at room temperature. The g -value of 2.013 was about the same as that observed for 2 (ref. 14).

Single crystals of $[2][I]$ were wired for four-point conductivity measurements along the stacking axis (x). Gold pads were first evaporated onto the crystals. Wires were attached to the pads with gold or silver paint, for experiments below or above room temperature, respectively. The conductivity (Fig. 4) is activated in the low temperature regime (< 200 K), and there are two points of inflection between 170 and 210 K; the low-temperature transition shows hysteresis (inset). From 210 K until decomposition the conductivity is weakly metallic. The highest conductivity is attained at ~ 350 K where $\sigma = 100$ S cm^{-1} . At a pressure of 1.5 GPa the conductivity is increased by a factor of about four, and the low-temperature phase transition is suppressed.

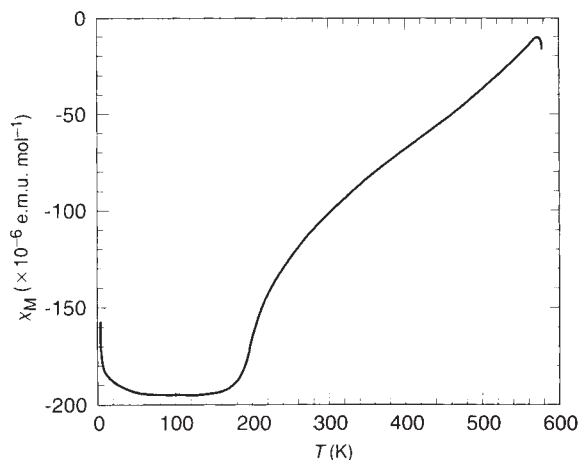


FIG. 3 The magnetic susceptibility of $[2][I]$ as a function of temperature. A Curie-Weiss fit to the low-temperature data gave a spin concentration of 0.04% on a per-molecule basis and the diamagnetism was found to be -197 p.p.m. e.m.u. mol^{-1} . Based on earlier solid-state magnetic susceptibility studies on undoped 2 (ref. 14) and using Pascal's constants the measured diamagnetism is consistent with the molecular formula.

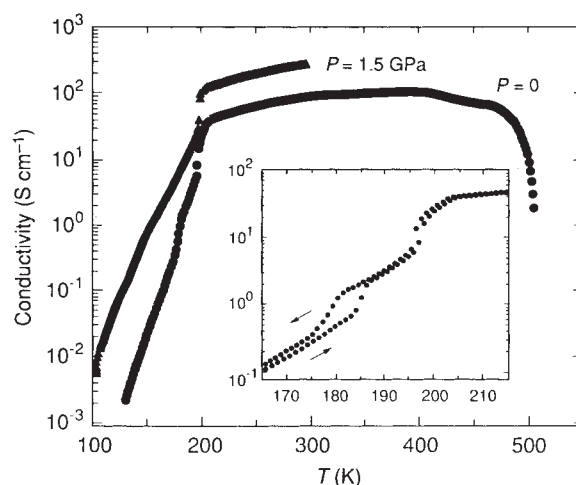


FIG. 4 The electrical conductivity of $[2][I]$ as a function of temperature and applied pressure (P).

Low-temperature X-ray diffraction was performed on a single crystal of $[2][I]$ using a Cu $K\alpha$ rotating-anode X-ray source and a closed-cycle helium refrigerator. A singly-bent pyrolytic graphite monochromator and a flat graphite analyser were used to increase the signal-to-noise ratio and minimize scattering from the cryostat Be windows. This arrangement results in a symmetric resolution function with a resolution width of ~ 0.01 Å $^{-1}$. Because of the large mosaic spread in crystals of $[2][I]$, scans in any direction but the radial direction are expected to show additional broadening from the mosaic structure of the crystal.

A search for a low temperature superlattice was undertaken by scanning parallel to $[h00]$, the chain axis, at various points

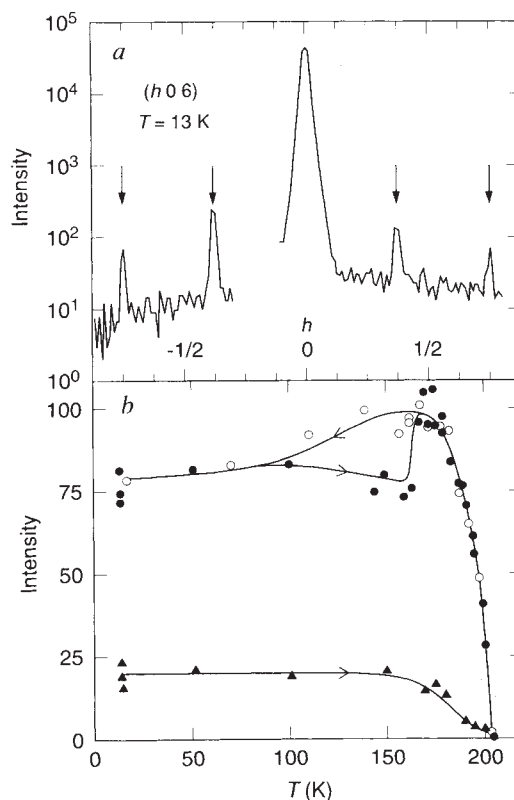


FIG. 5 a, Superlattice peaks observed for $[2][I]$ along $(h06)$ at 13 K. b, Temperature dependence of the superlattice reflections for $[2][I]$. Intensity is given in arbitrary units. Data for warming and cooling are shown by solid and open symbols respectively.

in reciprocal space. The results of scanning along $[h06]$ are shown in Fig. 5a. Small superlattice peaks marked with arrows can be seen at $h = \pm 3/8$ and $h = \pm 6/8$ around the (006) Bragg reflection. The third harmonic at $h = 9/8$ was not observed, nor was the Bragg reflection at (106), which is absent because of the body-centred lattice. Similar superlattice peaks were observed around nearly every major Bragg reflection in the $(0kl)$ zone, but were not observed around (00 l) Bragg reflections.

The temperature dependence of the (3/8 0 6) and the (6/8 0 6) superlattice reflections are shown in Fig. 5b. The onset transition is near 200 K in agreement with the magnetic susceptibility and conductivity measurements and is taken as evidence of the formation of a CDW. The increase of the intensity of the (3/8 0 6) reflection below the phase transition is quite steep, but apparently second order. The increase of the intensity of the second-order superlattice peak below the transition is less steep, as is expected because its intensity should be proportional to the fourth power of the order parameter rather than the second power.

There is an anomalous 'bump' in the intensity of the first-order satellite near 170 K on warming and a gradual decrease of the order parameter below 170 K on cooling. There is some ambiguity in determining the intensity of an X-ray from a scan through a broad, mosaic peak, however similar hysteretic behaviour, with a smaller range of hysteresis, is seen in the conductivity. This could be the signature of an additional CDW phase transition, but a superlattice from a second CDW has not been seen. The CDW wavevector is apparently commensurate at $h = 3/8$ at all temperatures in the range 10–300 K, however the low q -resolution of this experiment may have failed to detect a region of incommensurability near the phase transition.

In a one-dimensional conductor the CDW wavevector gives the band-filling directly¹⁶, and thus we infer three-eighth-full conduction band(s) in [2][I]. In our previous studies of neutral biradicals such as 2, the two dithiadiazolyl rings (1) were uncoupled and the biradicals could be treated as containing independent spin systems. If this argument is extended to [2][I] then every molecule of 2 is involved in two independent conduction bands each of which contains three-quarters of an electron per molecule for a net charge transfer of one-quarter of an electron per dithiadiazolyl ring to give $[2]^{1/2+}[I]^{1/2-}$. Alternatively, if there is a single conduction band which comprises both dithiadiazolyl rings, then the degree of charge transfer is one-half an electron on a per molecule basis from a conduction band which has a capacity of four electrons for the same overall division of charge. The iodines could consist of equal amounts of I^- and I_3^- .

The rise in the susceptibility is reminiscent of the behaviour shown by other neutral biradicals, particularly the 1,3-isomer of 2, which have shown giant paramagnetism¹³. The paramagnetism attained by [2][I] is much smaller and only corresponds to one-quarter of a spin per molecule on a Curie basis at the highest temperatures. The present case differs from that of the neutral biradicals in that the spins generated in [2][I] are itinerant and make a significant contribution to the conductivity. Interpreted as Pauli spins the highest value of the magnetic susceptibility corresponds to a density of states at the Fermi level, $n(E_F) \sim 6$ states (total) $(\text{eV mol})^{-1}$, considerably larger than would be expected for a one-dimensional metal based on the dithiadiazolyl system 1, near half-band filling. Furthermore the increase in magnetic susceptibility with temperature above 200 K and the flat conductivity suggests that carriers are still being generated at these temperatures.

The period of the CDW found in [2][I] is of note in view of our previous work on neutral radical-based molecular conductors^{11,13}. By definition the latter compounds give rise to a half-filled conduction band and in all cases we have found that the molecules are dimerized, except perhaps at high temperatures where magnetic excitations are detectable in some of the sulphur compounds¹³. Previously we have attempted to stabilize the met-

allic state and overcome the CDW in neutral radicals by using high pressure and by attempting to synthesize new systems with sufficient three-dimensional character to stabilize the Fermi surface against distortions. The present results suggest that doping the materials away from one-half to three-eighths band-filling also suppresses the formation of CDWs and is an effective route to conducting materials. The solid-state properties of [2][I] suggest that charge-transfer salts derived from neutral heterocyclic radicals represent a promising direction in the search for molecular electronic materials. □

Received 28 May; accepted 6 September 1993.

- Williams, J. M. et al. *Organic Superconductors* (Prentice-Hall, Englewood Cliffs, 1992).
- Haddon, R. C. et al. *Am. chem. Soc. Symp. Ser.* **481**, 71–89 (1992).
- Haddon, R. C. *Nature* **256**, 394–396 (1975).
- Soos, Z. G., Keller, H. J., Moroni, W. & Nothe, D. *Ann. N.Y. Acad. Sci.* **313**, 442–458 (1978).
- Nakasuji, K. et al. *J. Am. chem. Soc.* **111**, 9265–9267 (1989).
- Schultz, A. J. et al. *Inorg. Chem.* **26**, 3757–3761 (1987).
- Haddon, R. C. et al. *J. Am. chem. Soc.* **100**, 7629–7633 (1978).
- Haddon, R. C., Chichester, S. V., Stein, S. M., Marshall, J. H. & Muijsce, A. M. *J. Org. Chem.* **52**, 711–712 (1987).
- Oakley, R. T., Reed, R. W., Cordes, A. W., Craig, S. L. & Graham, J. B. *J. Am. chem. Soc.* **109**, 7745–7749 (1987).
- Hayes, P. J., Oakley, R. T., Cordes, A. W. & Pennington, W. T. *J. Am. chem. Soc.* **107**, 1346–1351 (1985).
- Cordes, A. W., Haddon, R. C. & Oakley, R. T. in *The Chemistry of Inorganic Ring Systems* (ed. Steudel, R.) 295–321 (Elsevier, Amsterdam, 1992).
- Banister, A. J. & Rawson, J. M. in *The Chemistry of Inorganic Ring Systems* (ed. Steudel, R.) 323–348 (Elsevier, Amsterdam, 1992).
- Andrews, M. P. et al. *J. Am. chem. Soc.* **113**, 3559–3568 (1991).
- Cordes, A. W. et al. *J. Am. chem. Soc.* **113**, 582–588 (1991).
- Liblong, S. W., Oakley, R. T. & Cordes, A. W. *Acta crystallogr. C* **46**, 140–142 (1990).
- Torrance, J. B. *Ann. N.Y. Acad. Sci.* **313**, 210–233 (1978).

ACKNOWLEDGEMENTS. We thank the NSERC (Canada) and the US NSF for support. C.D.B. acknowledges a DOE/ASTA traineeship.

Large contribution of organic aerosols to cloud-condensation-nuclei concentrations

T. Novakov* & J. E. Penner†

* Energy and Environment Division, Lawrence Berkeley Laboratory, One Cyclotron Road, Berkeley, California 94720, USA

† Global Climate Research Division, Lawrence Livermore National Laboratory, PO Box 808, Livermore, California 94551, USA

THE albedo and radiative properties of marine stratus clouds are determined largely by the number density of cloud condensation nuclei (CCN) over the oceans. Modelling studies have suggested that most of these nuclei are sulphate aerosols derived from both anthropogenic and natural sources¹. Here we present evidence that organic aerosols also play a key role in cloud nucleation. We determine the relative contributions of sulphate and organic aerosols to CCN concentrations at a marine site known to be influenced by anthropogenic emissions, and find that organic aerosols account for the major part of both the total aerosol number concentration and the CCN fraction. Thus, in regions that are affected by anthropogenic pollutants, organic aerosols may play at least as important a role as sulphate aerosols in determining the climate effect of clouds.

Our analysis is based on the results obtained in March–April 1992 as part of our field project on El Yunque peak (18° 19' N, 65° 45' W; elevation ~1,000 m), located on the eastern end of Puerto Rico. Previous results show that this site, exposed to trade winds, is subject to SO_4^{2-} concentrations ranging from ~300 to ~2,000 ng m^{-3} (ref. 3) and condensation nuclei concentrations ranging from ~250 to ~3,400 cm^{-3} . The occurrence of the higher concentrations is an indication that anthropogenic aerosols often strike the site. Not all aerosol particles (denoted

TABLE 1 Measured aerosol mass concentrations and derived number concentrations

Aerosol	Date	Mass (ng m^{-3})	N_{tot} (cm^{-3})	$N_{0.05}$ (cm^{-3})
Sulphate	30–31 March	954 (897–1,011)	196 (169–226)	162 (147–178)
	31 March–1 April	936 (878–994)	324 (288–359)	259 (237–279)
	7–8 April	984 (931–1,037)	203 (156–250)	95 (82–107)
	9–10 April	1,362 (1,280–1,444)	462 (417–505)	334 (308–359)
	Average	1,059 (996–1,121)	296 (257–335)	212 (193–230)
Organic	2–3 April	556 (415–697)	1,545 (950–2,130)	369 (251–487)
	4–5 April	896 (745–1,047)	1,147 (797–1,500)	413 (290–536)
	8–9 April	568 (442–694)	1,411 (865–1,953)	314 (206–422)
	10–11 April	628 (485–771)	1,827 (1,192–2,459)	371 (249–492)
	Average	662 (521–802)	1,482 (951–2,010)	366 (249–484)

Mass concentrations of sulphate as SO_4^{2-} and organic carbon as C. Numbers in parentheses are uncertainty ranges. N_{tot} is the total number concentration of particles with diameters $>0.03 \mu\text{m}$ and $N_{0.05}$ is the number concentration of particles with diameters $>0.05 \mu\text{m}$.

condensation nuclei (CN)) act as nuclei for the condensation of cloud droplets (that is, as CCN). Their ability to act as CCN depends on the size and chemical composition (water-soluble fraction) of these particles. Sulphate particles, owing to their chemical nature and size distribution, are known to be efficient CCN.

The data used in this discussion consist of (1) the mass size distributions of aerosol sulphur, chlorine, and organic carbon (OC), (2) CCN concentrations measured at 0.5% supersaturation, (3) CN concentrations, and (4) total non-sea salt (n.s.s.) SO_4^{2-} mass concentrations determined concurrently (using a co-located sampler) with CN and CCN measurements. The 24-h mass concentrations (sampling days for inorganic and organic analyses are shown in Table 1) of aerosol Cl, S and OC were determined on the eight stages of the micro-orifice uniform deposit impactor (MOUDI, MSP Corp.), having cut-off aerodynamic diameters D_{50} of 0.05, 0.10, 0.18, 0.32, 0.56, 1.0, 1.8 and $3.2 \mu\text{m}$.

The mass distributions measured by impactors are determined by the true aerosol mass size distribution and the collection efficiencies of the individual impactor stages. An approximation of the true size distribution can be obtained by using the Twomey inversion algorithm⁴ modified for multistage cascade impactors⁵ and the specific features of the MOUDI sampler. This data inversion procedure gives the mass distribution in a form suitable for numerical integration, and it extrapolates the mass size distribution to $0.025 \mu\text{m}$ aerodynamic diameter (one-half the cut-off size of the lowest impactor stage). The mass size distributions obtained in this way, after conversion from the measured aerodynamic to geometric particle diameters, were used to calculate the aerosol number size distributions and aerosol number concentrations. (The details of calculations and the analytical procedures are described in ref. 6). Inputs to the calculations consisted of aerosol OC (organic mass assumed to be $\text{OC} \times 1.2$ (ref. 7) with density of 1.0 g cm^{-3}), Cl (assumed NaCl with density of 2.165 g cm^{-3}), and S (assumed $(\text{NH}_4)_2\text{SO}_4$ with density of 1.769 g cm^{-3}) mass concentrations determined on the eight MOUDI stages.

Examples of mass size distributions of SO_4^{2-} , NaCl, and organic aerosols obtained by the numerical inversion of impactor data (collected on 9 and 10 April 1992) are shown in Fig. 1a. The derived aerosol number size distributions are shown in Fig. 1b. (The number size distribution, $n = dN/d \log D$, is related to the mass size distribution, $m = dM/d \log D$, through the relationship $n = (6/\pi D^3)(m/\rho)$, where ρ is the aerosol density, D is the particle geometric diameter and N and M are number and mass concentrations, respectively.) These data show that (1) most of the Cl (or sea salt) mass is associated with large particles ($D > 0.8 \mu\text{m}$); (2) the aerosol sulphur mass is confined to a relatively narrow range of smaller particles; (3) the OC size distribution is more evenly spread over the entire size range extending to the lowest size cuts, with a small fraction associated with the supermicrometre ($D > 1 \mu\text{m}$) particles; and (4) organic particles

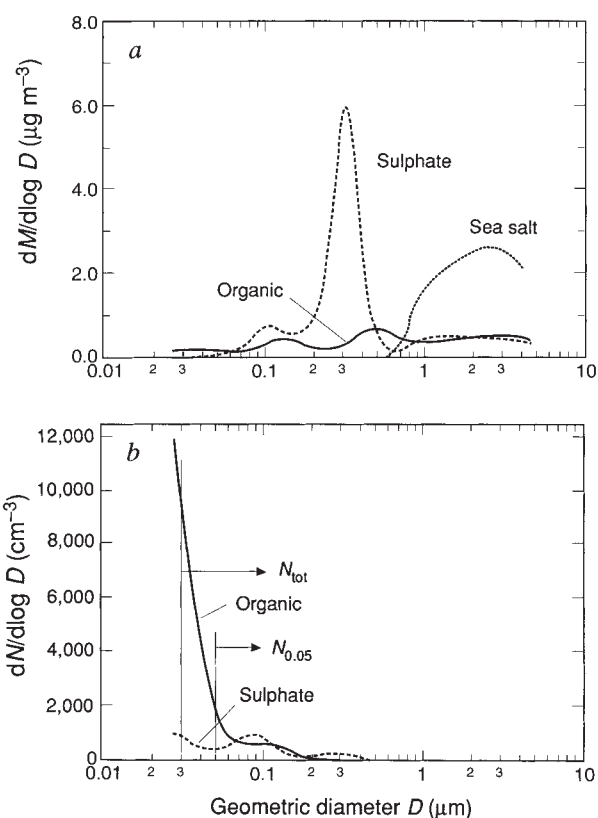


FIG. 1 a, Mass size distributions of sulphate, sea salt (NaCl), and organic aerosol measured on 9 April (inorganic) and 10 April (organic) 1992. b, Derived number size distributions. Vertical lines indicate the lower integration limits used to calculate N_{tot} and $N_{0.05}$ number concentrations.

contribute significantly more than sulphate aerosols to the number concentrations of small particles ($D < 0.08 \mu\text{m}$), despite the fact that the total SO_4^{2-} mass concentration is about two times higher than the OC mass. Because the Cl mass concentrations were below detection in small size impactor stages, the number size distributions could not be derived.

The aerosol number concentrations estimated from the derived number size distributions are listed in Table 1. (The uncertainty in derived values caused by the analytical errors in SO_4^{2-} and OC determinations are indicated by the numbers in parentheses and were derived by calculations using the upper and lower bounds of the analytical determinations.) Two sets of aerosol number concentrations, $N_{0.05}$ and N_{tot} , are given. $N_{0.05}$ refers to the number concentration of particles with diameters $>0.05 \mu\text{m}$. Ammonium sulphate particles in this size

TABLE 2 Measured n.s.s.* SO_4^{2-} , CCN and CN concentrations

Date	Non-sea-salt SO_4^{2-} (ng m^{-3})	CCN† (cm^{-3})	CN‡ (cm^{-3})	CCN/CN
28 March	1,116	589	1,533	0.38
29 March	1,241	407	1,677	0.24
30 March§	941	485	3,587	0.14
30 March	925	561	3,051	0.18
31 March	572	319	1,469	0.22
1 April	1,586	863	3,379	0.26
2 April	1,763	736	1,980	0.37
9 April	1,430	508	1,900	0.27
Measured average	1,197 (390)	558 (170)	2,322 (870)	0.26 (0.08)
Derived average		578	1,778	0.32

* Non-sea-salt.

† Measured with the M-1 counter (DH Associates).

‡ Measured with the CPC 3022 (TSI Corp.).

§ Morning data. All other data taken between 12:00 and 18:00 local time. Derived averages $\text{CCN} = N_{0.05}(\text{S}) + N_{0.05}(\text{O})$, $\text{CN} = N_{\text{tot}}(\text{S}) + N_{\text{tot}}(\text{C})$ from Table 1. Numbers in parentheses are standard deviations of the 6-h mean values.

range are expected to be activated at 0.5% supersaturation and therefore counted as CCN. For the reasons discussed below, we assume that the same criterion can be used to estimate the organic CCN concentration. N_{tot} refers to the total number concentration of particles with diameters $> 0.03 \mu\text{m}$. These values will be used for comparison with the CN concentrations measured at the site. The estimated N_{tot} concentrations, however, are expected to be lower than those measured because the CN counter used (CPC 3022, TSI Corp.) detects particles with diameters of $0.03 \mu\text{m}$ with 90% efficiency but also counts particles below this size, although with reduced efficiency.

Throughout the period when the MOUDI data were taken, clear air CN and CCN (at 0.5% supersaturation) concentrations were measured. Average 6-h (12:00–18:00 local time) CCN and CN number concentrations together with concurrently determined n.s.s. SO_4^{2-} mass concentrations are given in Table 2. The data show the mean n.s.s. SO_4^{2-} mass concentration of $1,197 \text{ ng m}^{-3}$ (range from ~ 600 to $\sim 1,800 \text{ ng m}^{-3}$), and the corresponding mean CCN and CN concentrations of 558 and $2,322 \text{ cm}^{-3}$, respectively. These measured values can be compared with the values of N_{tot} and $N_{0.05}$ (Table 1) estimated from SO_4^{2-} and OC mass size distributions.

The rationale for this comparison is based on the fact that the mean 6-h n.s.s. SO_4^{2-} concentration of $1,197 \text{ ng m}^{-3}$ (Table 2) is essentially the same as the mean MOUDI SO_4^{2-} concentration of $1,059 \text{ ng m}^{-3}$ (Table 1), and on the assumption that the mean OC concentration during the 6-h events was comparable to the mean of 662 ng m^{-3} obtained from impactor samples. The comparison of the data in Tables 1 and 2 shows that the average measured CCN concentration of 556 cm^{-3} is in remarkably good agreement with the sum of 578 cm^{-3} of average derived concentrations of $N_{0.05}(\text{S}) = 212$ and $N_{0.05}(\text{C}) = 366 \text{ cm}^{-3}$ (where C signifies organic carbon). The estimated N_{tot} concentrations should be considered as rough estimates (within $\sim 50\%$) because their derivation depends critically on the extrapolation of size distributions below the smallest impactor size cut. Nevertheless we note that the mean measured CN concentration of $2,322 \text{ cm}^{-3}$ is comparable to the sum of estimated $N_{\text{tot}}(\text{S})$ and $N_{\text{tot}}(\text{C})$ of $1,778 \text{ cm}^{-3}$ and that the mean measured CCN/CN ratio of 0.26 agrees approximately with the estimated $N_{0.05}/N_{\text{tot}}$ ratio of 0.32.

The average number concentrations were derived under the assumption that the organic and SO_4^{2-} components are found in separate particles (external mixture). Realistic aerosols, however, may consist of particles with mixed chemical composition (internal mixture). The uncertainty in the derived values caused by the lack of knowledge of the mixing state of the components can be evaluated by calculating the number concentrations for the aerosol with bulk composition equal to the average shown in Table 1, but assuming that the OC and SO_4^{2-} in each impactor stage are internally mixed and using the size-dependent volume

averaged density of the two components. The results show that the internal mixing decreases the derived total (sulphate + organic) $N_{0.05}$ number concentration by 8.4%, and the $N_{0.05}(\text{S})$ and $N_{0.05}(\text{C})$ concentrations by 6.9 and 9.2% respectively. Because the derivation of $N_{0.05}$ number concentrations does not depend critically on the assumed mixing state, the excellent agreement between measured CCN and derived total $N_{0.05}$ values strongly suggests that both sulphate and organic aerosol contribute to the CCN number concentration.

The apparent ability of organic aerosols to serve as CCN can have two possible, although not mutually exclusive explanations. The first is that the organic aerosols are intrinsically inactive and that condensation of sulphuric acid vapour, or some other water-soluble species, onto these particles renders them active CCN. This explanation is suggested by findings that aged atmospheric particles have a sulphuric acid coating⁸. Our finding that SO_4^{2-} constitutes $\sim 30\%$ of the mass in the two lowest-size impactor stages tends to support this view. The second explanation is that a fraction of the chemically complex organic aerosol material is hygrophilic. Circumstantial evidence in support of this possibility is provided by the known water-soluble nature of the ambient organic aerosol^{9,10}. The present data, however, are insufficient to argue for either of these possibilities.

The results presented above show that $\sim 37\%$ of the CCN number concentrations measured at the site can be accounted for by sulphate, and the remaining 63% by the organic aerosol mass concentration. Obviously, if the organic aerosol is intrinsically CCN active, the production of these particles will determine the organic CCN concentration. If, however, the condensation of sulphuric acid vapour renders the organic particles active, then the number concentration and surface area of these particles, together with existing sulphate particles, will control the total number of CCN.

Anthropogenic SO_4^{2-} mass is believed to be produced principally by oxidation of SO_2 in cloud droplets. This formation pathway would not produce new sulphate CCN, only larger particles formed on pre-existing aerosols^{11,12}. In contrast, the formation of organic aerosol does not involve aqueous chemistry. It is either derived from direct emissions of primary particles or produced by condensation of low-vapour-pressure products of gas-phase reactions of reactive hydrocarbons. Therefore, it can be expected that the relationship between aerosol number concentration and mass emission rates should be stronger for organic species than for sulphur species. The same should be expected for the fraction of organic aerosols that serve as CCN, regardless of whether their nucleation properties are intrinsic to the organic compounds or acquired by adsorption of inorganic species, such as sulphuric acid.

Our estimates, taken together with the fact that OC concentrations at our site are within the range of OC concentrations measured in other marine locations^{13–15} and with inferences from

isotopic studies indicating that the sub-micrometre organic aerosol is primarily of continental origin¹⁶, suggest that organic matter may be a major contributor to CCN in the anthropogenically influenced Northern Hemispheric marine atmosphere. □

Received 8 April; accepted 9 September 1993.

1. Charlson, R. J. et al. *Science* **255**, 423–430 (1992).
2. Penner, J. E., Dickinson, R. & O'Neill, C. *Science* **256**, 1432–1434 (1992).
3. Novakov, T., Rivera-Carpio, C., Penner, J. E. & Rogers, C. F. *Tellus* (in the press).
4. Twomey, S. J. *Comput. Phys.* **18**, 188–200 (1975).
5. Winklmayr, W., Wang, H. C. & John, W. *Aerosol Sci. Technol.* **13**, 322–331 (1990).
6. Rivera-Carpio, C., Novakov, T., Chow, J. C. & Rogers, C. F. Report No. LBL-33905 (Lawrence Berkeley Lab., Berkeley, 1993).
7. Gray, H. A., Cass, G. R., Huntzicker, J. J., Heyerdahl, E. K. & Rau, J. A. *Sci. tot. Envir.* **36**, 17–25 (1984).
8. Bigg, E. K. J. *Appl. Met.* **19**, 521–533 (1980).
9. Cadle, S. H. & Groblicki, P. J. in *Particulate Carbon-Atmospheric Life Cycle* (eds Wolff, G. T. & Klimisch, R. L.) 89–109 (Plenum, New York, 1982).
10. Mueller, P. K., Fung, K. K., Heisler, S. L., Grosjean, D. & Hidy, G. M. in *Particulate Carbon Atmospheric Life Cycle* (eds Wolff, G. T. & Klimisch, R. L.) 343–368 (Plenum, New York, 1982).
11. Langner, J., Rodhe, H., Crutzen, P. J. & Zimmerman, P. *Nature* **359**, 712–716 (1992).
12. Lelieveld, J. & Heintzenberg, J. *Science* **258**, 117–120 (1992).
13. Hoffman, E. J. & Duce, R. A. *Geophys. Res. Lett.* **4**, 449–452 (1977).
14. Eichmann, R. et al. *Atmos. Envir.* **13**, 587–599 (1979).
15. Wolff, G. T. et al. *Atmos. Envir.* **20**, 1229–1239 (1986).
16. Cachier, H., Buat-Ménard, P., Fontugne, M. & Rancher, J. *Tellus* **38B**, 161–177 (1986).

ACKNOWLEDGMENTS. This work was supported by the US Department of Energy. We thank J. Ogren and E. K. Bigg for comments and suggestions, and C. Rivera-Carpio for performing the calculations.

Cumulates from strongly depleted mid-ocean-ridge basalt

Kent Ross*† & Don Elthon†‡

* Department of Geosciences, † Department of Chemistry, ‡ Texas Center for Superconductivity, University of Houston, Houston, Texas 77204, USA

RECENT studies of abyssal peridotites¹, mid-ocean-ridge basalts (MORBs)² and their entrained melt inclusions^{3,4} have shown that fractional melting of the upwelling sub-oceanic mantle produces magmas with a much wider range of compositions than erupted MORBs. In particular, it seems that strongly depleted primary magmas are routinely produced by melting beneath ridges¹. The absence of strongly depleted melts as erupted lavas prompts the question of how long such magmas survive beneath ridges, before their distinctive compositions are concealed by mixing with more enriched magmas. Here we report mineral compositions from a unique suite of oceanic cumulates recovered from DSDP Site 334 (ref. 5), which indicate that the rocks crystallized from basaltic liquids that were strongly depleted in Na, Ti, Zr, Y, Sr and rare-earth elements relative to any erupted MORB. It thus appears that the magmatic plumbing system beneath the Mid-Atlantic Ridge permitted strongly depleted magmas to accumulate in a magma chamber and remain sufficiently isolated to produce cumulate rocks. Even so, spatial heterogeneity in the compositions of high-calcium pyroxenes suggests that in the later stages of solidification these rocks reacted with infiltrating enriched basaltic liquids.

Deep Sea Drilling Project Site 334 was drilled 104 km west of the Mid-Atlantic Ridge at 37° 02' N, 34° 25' W. Plagioclase lherzolite, olivine gabbro, gabbro and noritic anorthosite were recovered below a thin cover of Miocene and younger sediments and MORBs⁶. Petrological studies of the cumulate rocks have shown that they contain olivine (Fo_{90–85}), high-Ca and low-Ca pyroxenes (Mg number, Mg# = 100 Mg/(Mg + Fe) = 91–70) and plagioclase (An_{90–75})^{6–8}. Compared to other suites of oceanic cumulates^{9–13}, Site 334 cumulates have anomalously high anorthite in plagioclase, and low Na and Ti in high-Ca pyroxene (Fig. 1a–c). Experimental and phenocryst studies of erupted MORBs show that they crystallize miner-

als similar to other oceanic cumulates but unlike Site 334 cumulates¹⁴.

Site 334 peridotites are thought to be cumulate, not residual, based on their textures, modes and compositions. These rocks have classic cumulate textures. The samples are not strongly tectonized, and high-Ca pyroxene (clinopyroxene, cpx) modal abundances exceed that of low-Ca pyroxene (orthopyroxene, opx) in peridotite. It has been suggested that the peridotites from Site 334 are melt-impregnated residual mantle¹⁵; however, the similarity in mineral compositions between rocks that are unquestionably cumulate (gabbroic rocks) and the peridotites shows that the peridotites are part of the same cumulate fractionation series. Plagioclase is interstitial, but its high anorthite content (An_{88–90}) indicates that it does not represent simple *in situ* crystallization of trapped melt. Unambiguous assignment of the pyroxenes in peridotites as cumulus or post-cumulus is not possible. In many samples, pyroxene textures could be interpreted as post-cumulus, but the geochemical characteristics of pyroxenes indicate that they are cumulus or heteradecumulus (uniform and high Mg# and high Cr abundances, very low incompatible element abundances). Clinopyroxene grains are large (3–8 mm) and have complicated shapes, sometimes enclosing olivine and spinel. Orthopyroxene is an abundant cumulus phase in olivine gabbro and gabbro. Abundant, early crystallization of magnesian opx suggests that parental magmas of Site 334 cumulates were high SiO₂ (52–55 wt %) liquids derived by relatively low pressure, late-stage melting of strongly depleted abyssal peridotite.

The compositions of cpx in three peridotites and one olivine gabbro are presented in Table 1. Similar data for three

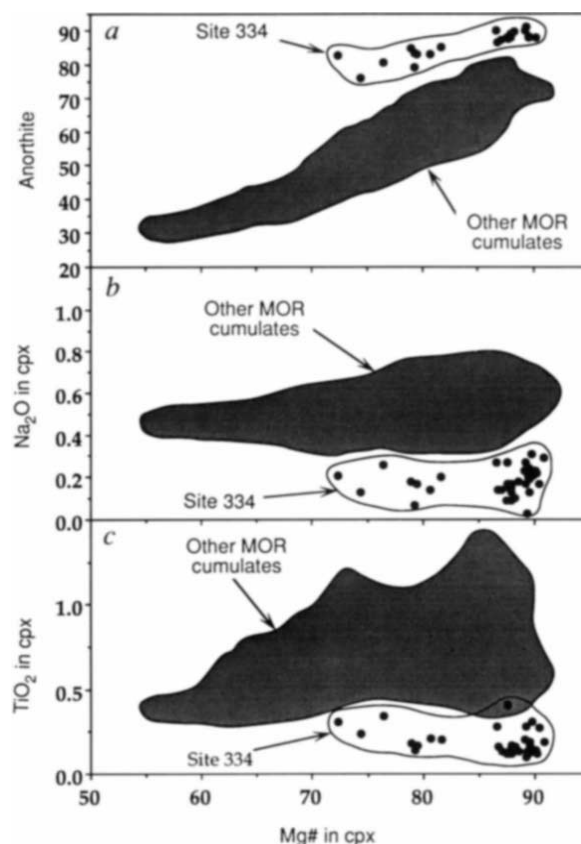


FIG. 1 Major and minor element mineral compositions in oceanic cumulate rocks determined by electron microprobe. a, Anorthite in plagioclase plotted against Mg number (see text) in clinopyroxene (cpx) b, Na₂O in cpx plotted against Mg number in cpx. c, TiO₂ in cpx plotted against Mg number in cpx. Data for Site 334 are from refs 7–9 and K.R. and D.E., unpublished data. Data for other mid-ocean ridge (MOR) cumulates are from refs 10–14.

other peridotites are included in Fig. 2a and b. These data were obtained on the CAMECA 3f ion probe at Woods Hole Oceanographic Institution. Three to four spots in each sample on one or more grains were analysed. Individual analyses are presented rather than average compositions because the data vary within each sample by significantly more than analytical uncertainties.

Rare earth element (REE) compositions of cpx at Site 334 and in abyssal peridotites¹ are shown in Fig. 2a and b. Abyssal peridotite data are from the Vulcan Fracture Zone. Other, more depleted cpxs in abyssal peridotites are near or at hot spots. Site 334 data are shown in groups of depleted cpx (Fig. 2a, La_n is $0.10\text{--}0.30 \times$ chondrites (where subscript n indicates abundances are normalized to chondritic meteorite abundances) and enriched cpx (Fig. 2b, La_n is $0.30\text{--}1.0 \times$ chondrites). Both enriched and depleted cpx compositions were found in all samples and within some individual grains. Analyses of REE in bulk cpx mineral separates from this suite by instrumental neutron activation analysis (INAA) have a range of compositions very similar to those determined in the ion-probe analyses. Depleted cpx more accurately reflects original compositions, whereas enriched cpx is thought to reflect infiltration metasomatism.

The estimated REE compositions of magmas that crystallized Site 334 depleted cpx are shown in Fig. 2c (calculated using cpx/liquid partition coefficients of ref. 16) along with REE patterns of representative normal MORB (N-MORB) and Site 334 erupted basalts. Table 2 shows estimates of the compositions of the most depleted Site 334 parental liquids, N-type MORBs, and erupted Site 334 basalts. Estimated parental liquids of Site 334 cumulates were depleted in Na by a factor of ~ 1.5 , Y by a factor of ~ 2 , Sr and Ti by a factor of ~ 3 , and Zr by a factor of ~ 4 relative to the lowest abundances in N-type MORBs. REE depletion factors range from ~ 2 at Yb to ~ 4 at the light REEs (LREE). Magmas parental to these cumulates were strongly depleted relative to N-type MORBs and to Site 334 erupted basalts. The precise values of these depletion factors are, of course, dependent upon the exact partition coefficients (D_s) used to calculate the inferred liquid compositions. Using different D_s would change the estimated levels of depletion, but would not change the overall conclusion that the parental liquids of these rocks were strongly depleted relative to erupted MORBs (see Table 2 for D_s used to calculate parental liquids and caption for Fig. 2 for further discussion of D_s).

Site 334 depleted cpxs have lower Nd to Yb abundances than Vulcan Fracture Zone abyssal peridotite cpx, but are not lower in La and Ce. Site 334 inferred liquids do not, therefore, appear to be as strongly LREE-depleted as some fractional melts inferred from abyssal peridotites¹ and melt inclusions⁴. It is unclear whether the relatively flat inferred liquids for Site 334 are correct estimates of the LREE liquid composition. We cannot rule out the possibility that the metasomatism proposed below has altered even the most depleted cpxs in these rocks.

Individual rocks have highly variable LREE concentrations in cpx analysed by ion probe. One cpx grain in sample 24-4 7-13, for example, has both the most LREE-depleted (La, 0.04 p.p.m.; Ce, 0.20 p.p.m.) and the most LREE-enriched (La, 0.31 p.p.m.; Ce, 0.91 p.p.m.) composition. In Table 1, spots taken on one grain in sample 24-4 7-13 are labelled according to their position on the grain. The 'c' indicates that this spot was nearest the exposed centre of the grain, the 'i' stands for intermediate and indicates that these spots were relatively closer to the rim. The spot nearest the core is most depleted and those nearer the rim are relatively enriched, consistent with reaction of depleted cpx with infiltrating enriched magmas.

Clinopyroxene $(La/Sm)_n$, $(La/Ce)_n$, $(La/Nd)_n$ and $(La/Yb)_n$ ratios (where subscript n indicates abundances are normalized to chondritic meteorite abundances) are highly variable and exhibit trends with La abundance. Figure 2d shows that $(La/Ce)_n$ increases steeply with increasing La_n , reaches a maximum and then declines with further increase in La_n . The $(La/Nd)_n$ ratio

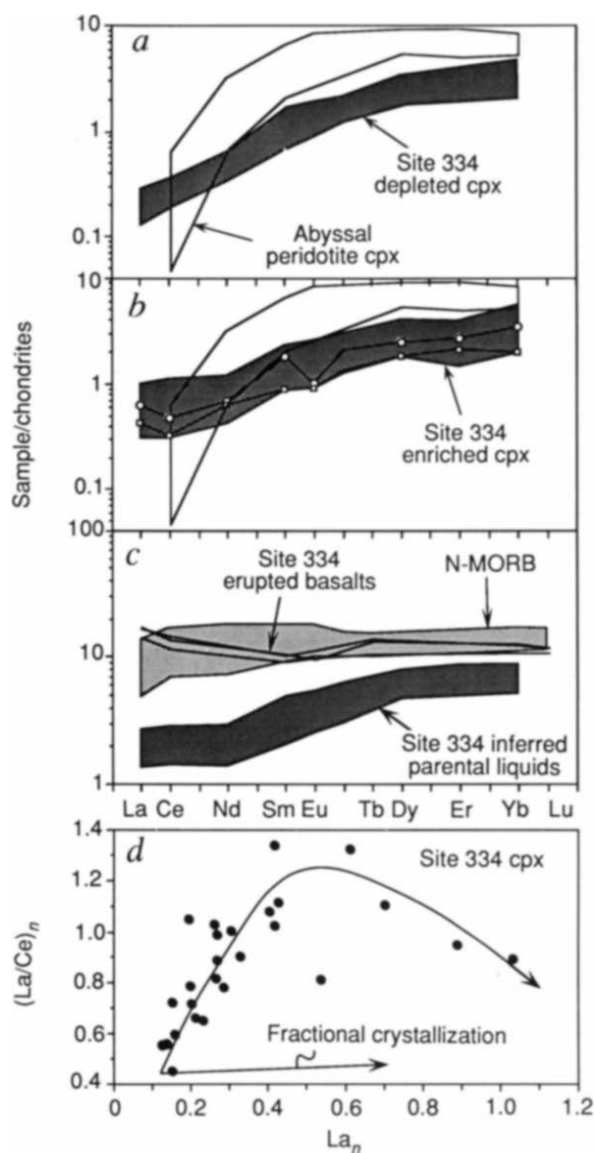


FIG. 2 Rare earth element (REE) compositions of clinopyroxene (cpx) in Site 334 cumulates and abyssal peridotites and in inferred liquids in equilibrium with Site 334 cumulates. a, REE abundances in depleted cpx in Site 334 cumulate peridotites and olivine gabbroanorite. This group of analyses had the lowest light REE (LREE) abundances ($La_n < 0.30$) and absent or minor LREE kinks in their patterns relative to the enriched cpx shown in b. b, REE abundances in enriched cpx in Site 334 cumulates. Two examples of LREE kinked patterns are shown. c, Inferred liquids that crystallized Site 334 depleted cpx are compared to compositions of a representative group of N-MORBs²⁴⁻²⁶ and to Site 334 erupted basalts. Liquids in equilibrium with cpx were calculated using the REE cpx/liquid partition coefficients of ref. 16 (set B). Some REE partitioning studies indicate lower cpx/liquid partition coefficients^{21,28} at the LREEs than we have used, and this would shift the field for equilibrium liquids at La and Ce upward on this figure. However, the inferred liquids that crystallized Site 334 cumulates are depleted relative to N-MORB, using partition coefficients from a wide variety of studies^{21,23,28,29}. The set we have used in plotting this figure (listed in the legend to Table 2) falls near the centre of the range of partition coefficients determined from a variety of studies. d, La_n plotted against the $(La/Ce)_n$ ratio in ion-probed cpx grains. The trend labelled fractional crystallization would be produced by both fractional crystallization and by overgrowths of cpx from evolving intercumulus liquids on the original cumulus cpx cores. The observed trend indicates chromatographic metasomatism (further discussion in text). All REE data displayed in these diagrams are normalized to the chondrite values in ref. 30 multiplied by 1.27.

TABLE 1 Compositions of clinopyroxenes

Sample	Grain	Mg #	La	Ce	Nd	Sm	Eu	Dy	Er	Yb	Ti	V	Cr	Sr	Y	Zr
334 22-2 41-43 peridotite	1	89.6	0.13	0.32	0.51	0.43	0.14	0.99	0.71	0.92						
	1		0.06	0.25	0.30	0.32	0.09	0.68	0.46	0.63	633	241	7856	2.2	4.9	1.5
	1		0.13	0.25	0.36	0.17	0.06	0.57	0.42	0.40	556	220	7903	1.7	3.7	1.2
	1		0.06	0.15	0.21	0.14	0.08	0.55	0.41	0.42	534	219	8074	1.7	3.9	0.9
334 26-2 5-10 peridotite	1	90.9	0.06	0.22	0.35	0.20	0.14	0.65	0.50	0.55	651	229	6826	2.7	4.6	0.9
	1		0.05	0.17	0.24	0.23	0.08	0.68	0.66	0.51	764	238	7118	1.5	4.6	1.3
	2		0.12	0.30	0.39	0.31	0.11	0.71	0.53	0.65						
	3		0.05	0.21	0.20	0.19	0.08	0.61	0.65	0.61	677	238	6832	2.3	3.8	1.5
334 24-4 7-13 olivine gabbro-norite	c	87.9	0.04	0.20	0.38	0.18	0.13	0.65	0.56	0.61	662	235	6638	2.2	4.4	1.7
	i		0.31	0.91	0.66	0.44	0.13	0.92	0.69	0.76	857	264	8731	3.1	5.6	1.6
	i		0.08	0.26	0.36	0.21	0.10	0.70	0.30	0.44	531	224	4522	1.7	3.2	1.1
	i		0.13	0.30	0.36	0.30	0.11	0.72	0.52	0.59						
334 22-2 83-88 peridotite	1	89.1	0.08	0.24	0.35	0.30	0.08	0.75	0.62	0.50	539	222	7703	1.7	4.9	1.4
	2		0.16	0.52	0.68	0.44	0.18	1.27	0.80	1.12						
	3		0.27	0.74	0.55	0.27	0.09	0.86	0.77	0.67	854	231	7594	9.7	5.6	1.5
	3		0.05	0.27	0.36	0.22	0.13	0.78	0.57	0.56	620	225	7749	1.8	5.6	1.5

All trace element data in p.p.m. Analytical uncertainties: Er, Yb, Sr, Zr, 12–15%; Ce, Nd, Sm, Eu, 20%; La 35%; Ti, V, Cr 1.5–2%; Y 5%. Mean Mg # in cpx are averages of 8–10 microprobe analyses per sample. Grain designations indicate number of grains analysed in each sample. c indicates analysis nearest core, i indicates analysis at an intermediate location between the centre and the rim of the grain. Analytical techniques are described in ref. 1.

increases and then reaches a plateau, whereas $(La/Yb)_n$ steadily increases with increasing La_n . These trends are precisely like those predicted in modelling of the geochemical effects of chromatographic metasomatism^{17,19} by infiltrating, enriched basaltic liquids; we accordingly suggest that cpx grains in these rocks have been altered by this process. (The resulting enrichment in light rare-earth elements produces cpx REE patterns with abrupt changes in slope at Ce or Nd, such as those shown in Fig. 2b.) These observations rule out simple post-cumulus overgrowth from evolving interstitial liquids as the cause of the LREE variability, as this would produce trends of La versus La/REE ratios with a near-zero slope (trend labelled fractional crystallization in Fig. 2d). Other evidence that LREE variability results from metasomatism of depleted grains rather than from liquid trapping and post-cumulus overgrowth is found in the very high Cr content of LREE-enriched cpx (see Table 1). Grain growth from evolving intercumulus liquid would produce a trend of sharply declining Cr with increasing REE abundances.

Initial crystallization and accumulation of the cumulus mineral grains were followed by interaction with later more enriched liquids that locally produced LREE-enriched pyroxenes. The composition of the initial infiltrating melt is not known in detail, and cannot be directly calculated from the compositions of enriched grains obtained thus far. Chromatographic metasomatism produces increases in La/REE ratios in metasomatized grains above that which would be in equilibrium with the initial infiltrating melt. LREE enrichment in cpx indicates that the infiltrating melt must be a LREE-enriched basalt, quite different from the depleted basalts from which these cumulates initially crystallized. Infiltration of depleted cumulates by enriched basaltic liquids could have taken place by porous flow in a crystal mush zone at deep levels in the oceanic crust, such as that proposed on the basis of seismic data from the East Pacific Rise²⁰.

Previous evidence for strongly depleted MORBs^{1,4} has established the existence of such melts, but did not provide information about where these melts mix with more enriched MORB liquids. Although this suite of cumulates provides supporting evidence for strongly depleted liquids, the most important result derived from this study is that it shows that such magmas have collected in crustal level magma chambers and remained sufficiently isolated to form these distinctive cumulates. Thus, at least in some cases, aggregation and mixing of compositionally distinct, small-degree fractional melts takes place in the crust, and not during magma ascent in the upper mantle. Interaction between depleted cumulates and more enriched liquids suggests that new melt could be delivered into mid-ocean-ridge magma chambers in some cases by porous flow through pre-existing cumulates rather than by direct injection through dikes. □

TABLE 2 Estimated Site 334 most-depleted liquids

	Site 334 liquids estimated from most-depleted cpx	N-MORB erupted basalts
Na ₂ O (wt%)	0.90	1.5
TiO ₂ (wt%)	0.20	0.60
Sr (p.p.m.)	15	50–150
Zr (p.p.m.)	6	25–150
Y (p.p.m.)	7	15–30
La _n	1.3	5–15
Ce _n	1.6	6–17
Nd _n	1.7	7–18
Sm _n	2.4	9–20
Eu _n	2.5	10–19
Dy _n	4.9	13–20
Er _n	5.4	13–20
Yb _n	5.0	11–19

Estimated Site 334 liquids are calculated from most-depleted cpx analyses, using the following cpx/liquid distribution coefficients: Na₂O, 0.15; TiO₂, 0.40; Sr, 0.10; Y, 0.45; Zr, 0.15; La, 0.095; Ce, 0.12; Nd, 0.24; Sm, 0.32; Eu, 0.36; Dy, 0.37; Er, 0.37; Yb, 0.40. REE partition coefficients are from ref. 16 (set B). Other values are representative of distribution coefficients from refs 21–23. Sources for N-MORBs are refs 24–26. Data for Site 334 erupted basalts are from refs 6, 27 and K.R. and D.E., unpublished data.

Received 4 June; accepted 3 September 1993.

- Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. *J. geophys. Res.* **95**, 2661–2678 (1990).
- McKenzie, D. *Earth planet. Sci. Lett.* **72**, 149–157 (1985).
- Sobolev, N. V. & Shimizu, N. *Nature* **363**, 151–154 (1993).
- Hummel, E. & Whitechurch, H. *Earth planet. Sci. Lett.* **88**, 169–181 (1988).
- Aumento, F. et al. *Init. Rep. DSDP Leg 37*, 239–273 (1977).
- Hodges, F. N. & Papke, J. J. *J. geophys. Res.* **81**, 4135–4151 (1976).
- Clarke, D. B. & Loubat, H. *Init. Rep. DSDP Leg 37*, 847–856 (1977).
- Symes, R. F., Bevan, J. C. & Hutchison, R. *Init. Rep. DSDP Leg 37*, 841–845 (1977).
- Tiezzi, L. J. & Scott, R. B. *J. geophys. Res.* **85**, 5438–5454 (1980).
- Meyer, P. S., Dick, H. J. B. & Thompson, G. *Contr. Miner. Petrol.* **103**, 44–63 (1989).
- Bloomer, S. H., Natland, J. H. & Fisher, R. L. in *Magmatism in the Ocean Basins, Spec. Publ. No. 42* (eds Saunders, A. D. & Norry, M. J.), 107–124 (Geol. Soc. Lond., 1989).
- Elthon, D. *J. geophys. Res.* **92**, 658–682 (1987).
- Hebert, R., Constantin, M. & Robinson, P. T. *Proc. ODP Sci. Res.* **118**, 3–20 (1991).
- Elthon, D., Stewart, M. & Ross, D. K. *J. geophys. Res.* **97**, 15189–15200 (1992).
- Girardeau, J. & Francheteau, J. *Earth planet. Sci. Lett.* **115**, 137–149 (1993).

16. Grutzeck, M. W., Kridelbaugh, S. J. & Weill, D. F. *Geophys. Res. Lett.* **1**, 273–275 (1974).
17. Navon, O. & Stolper, E. *J. Geol.* **95**, 285–307 (1987).
18. Takazawa, E., Frey, F. A., Shimizu, N., Obata, M. & Bodinier, J. L. *Nature* **359**, 55–58 (1992).
19. Bodinier, J. L., Vasseur, G., Vernieres, J., Dupuy, C. & Fabries, J. J. *Petrol.* **31**, 597–628 (1990).
20. Sinton, J. M. & Detrick, R. S. *J. geophys. Res.* **97**, 197–216 (1992).
21. Hart, S. R. & Dunn, T. *Contr. Miner. Petrol.* **113**, 1–8 (1993).
22. Johnson, K. T. M. & Kinzler, R. J. *EOS* **70**, 1388 (1989).
23. Dunn, T. *Contr. Miner. Petrol.* **96**, 476–484 (1987).
24. LeRoex, A. P. et al. *J. Petrol.* **97**, 197–216 (1992).
25. Frey, F. A., Bryan, W. B. & Thompson, G. J. *geophys. Res.* **79**, 5507–5525 (1974).
26. Bender, J. F., Langmuir, C. H. & Hanson, G. N. *J. Petrol.* **25**, 213–254 (1984).
27. Dostal, J. & Mueke, G. K. *Init. Rep. DSDP Leg 37*, 574–575 (1977).
28. McKay, G. A., Wagstaff, J. & Yang, S.-R. *Geochim. cosmochim. Acta* **50**, 927–937 (1986).
29. Fujimaki, H., Mitsunobu, T. & Aoki, K. *J. geophys. Res.* **89**, B662–B672 (1984).
30. Anders, E. & Ebihara, M. *Geochim. cosmochim. Acta* **46**, 2363–2380 (1982).

ACKNOWLEDGEMENTS. This work has been supported by the US National Science Foundation and the Texas Advanced Research Program. We thank N. Shimizu and D. Mittlefehdt for assistance with the analytical program, and N. Shimizu and F. A. Frey for thoughtful reviews of the manuscript.

Persistent patterns in the geomagnetic field over the past 2.5 Myr

David Gubbins & Peter Kelly

Department of Earth Sciences, Leeds University, Leeds LS2 9JT, UK

HISTORICAL geomagnetic measurements covering the past 400 years reveal a symmetrical pattern of four relatively stationary flux concentrations ('lobes') at the surface of the liquid core and regions of rapid change extending from the Atlantic to the Indian Ocean¹. Palaeomagnetic data define a time-average over several thousand years² which might reflect the stationary parts of the present field, but unfortunately the historical record is too short

to provide a satisfactory average³. Here we model palaeomagnetic directions from the past 2.5 Myr using the same methods as for modern data⁴, and find the two northern lobes in the same position as today, over Arctic Canada and Siberia. In southern regions, by contrast, the field appears to have been smoothed out, as might be expected from the current rapid secular variation¹. We propose that the present geomagnetic field morphology and pattern of secular variation have persisted for several million years, as would occur if the solid mantle controls flow at the top of the core.

There are two major data sets available for palaeomagnetic directions: one of declinations and inclinations derived from lavas⁵ and one of inclinations derived from marine sediments⁶. We obtained the former from P. McFadden (ref. 5), and the latter from D. Kent (personal communication). We then selected measurements made from the Brunhes and Matuyama polarity chrons and separated normal and reversed polarities (Fig. 1). Note that a substantial number of normal data are available

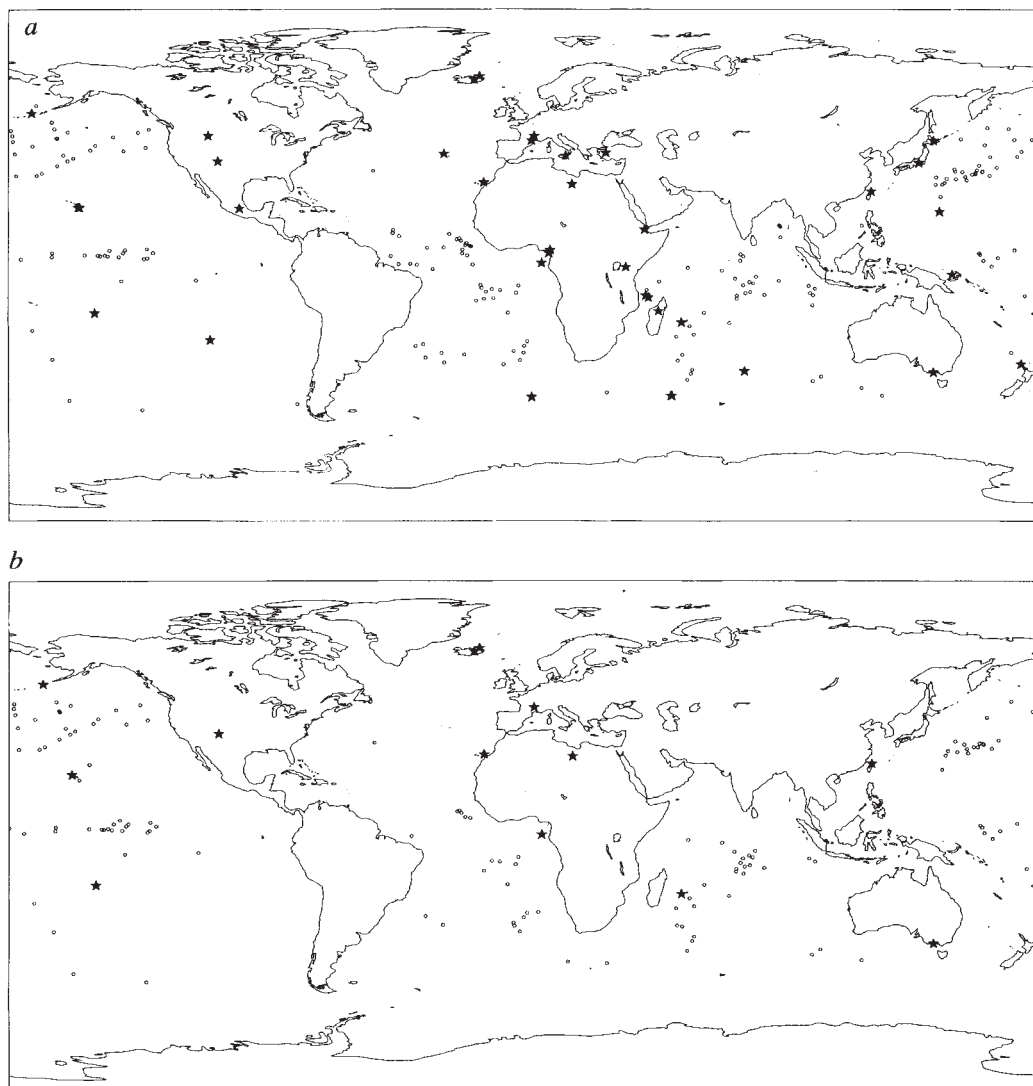


FIG. 1 Sites of normal (a) and reversed (b) polarity, showing sites of lava data (stars) and sites of marine sediment data (circles).

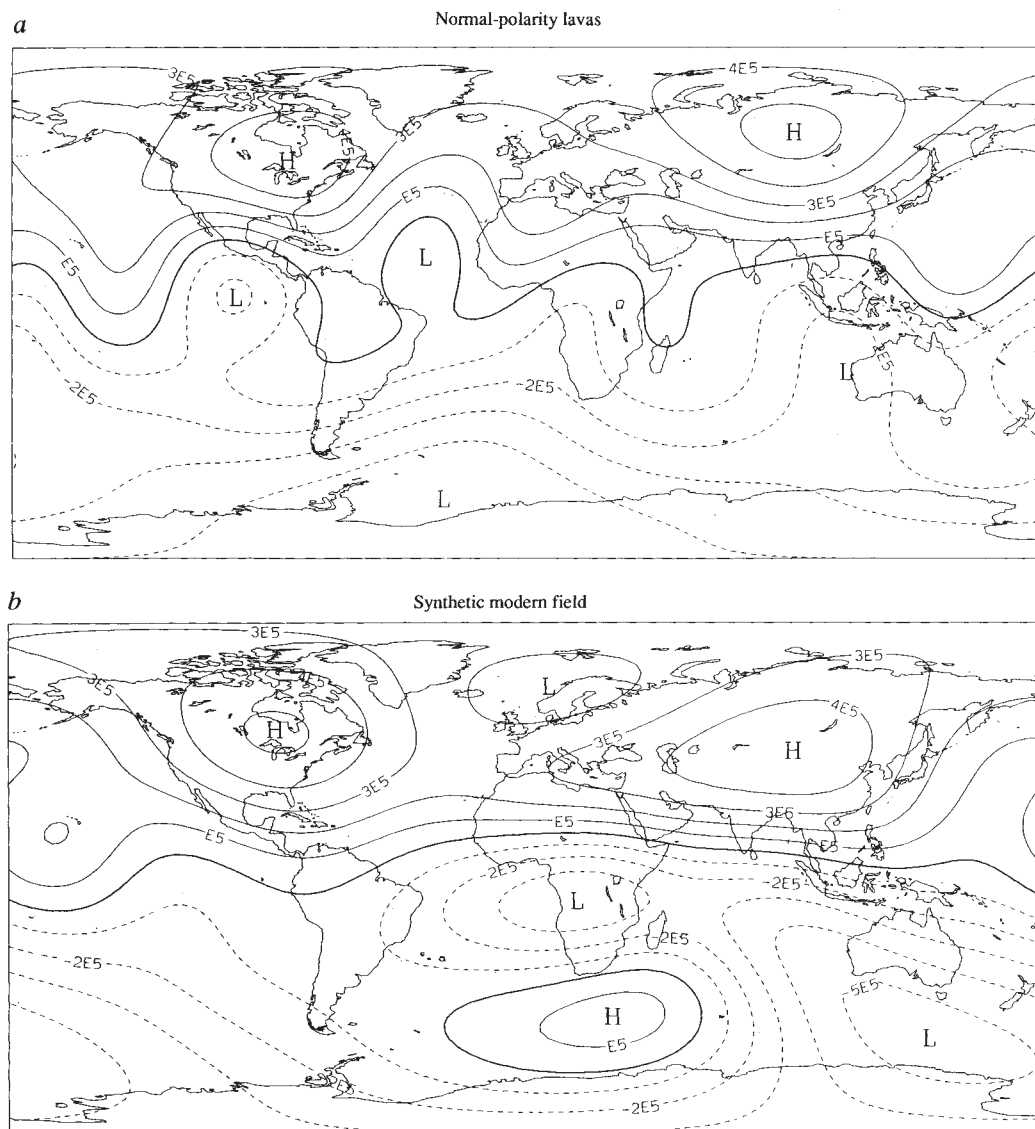
from short reversals within the reversed Matuyama chron. We ignore the effect of plate movements on the site positions during this 2.5-Myr interval. Each site represents the average of a number of measurements, which is thought to remove the effects of secular variation⁷. This averaging reduces the volume of lava data from 1,247 individual samples to 53 average values of declination and inclination. The standard deviation of this mean ($\sigma \approx 1^\circ$ for both data sets) gives an estimate of the internal consistency of the data from one site. Note that in Fig. 1 the sites are often clustered together; nearby sites give similar directions for lavas but for marine sediments the directions differ by about 4° or 4σ , much greater than expected from the present-day field. We attribute the difference to unknown systematic errors at individual sites, possibly associated with the coring process or local effects on the magnetization of the sediments and therefore increased σ (to 4°) for the sediment data. As an additional precaution we analysed lavas and marine sediments separately to check consistency. During the course of this work we obtained additional data from X. Quidelleur and J-P. Valet (personal communication), who also recommended removal of six sites from the McFadden data set. These changes do not alter our conclusions, and will be included in a fuller description elsewhere (P.K. and D.G., manuscript in preparation).

First, data for normal-polarity lavas from both the Brunhes and Matuyama chrons were modelled to 1σ for a smooth core field using the methods developed for modern and historical data

described in full elsewhere⁴. The iterative procedure starts from an axial dipole field to avoid any bias from the starting field. The trace of the resolution matrix, often interpreted as the number of degrees of freedom in the model, was 26. The radial component of magnetic field at the core-mantle boundary is shown in Fig. 2a. It is very much smoother than similar maps of the modern field; this may arise from the time averaging or because of fewer data. To assess the latter effect we substituted modern (1980) data in the palaeomagnetic sites and subjected it to the same analysis (Fig. 2b). The method smooths the model to fit the sparse data set so Fig. 2b contains less detail than is typical for modern data, but it gives an idea of the types of feature that may be resolved with the palaeomagnetic data set. The resemblance to the palaeofield in the Northern Hemisphere is striking because it exhibits the familiar two flux lobes over Canada and Siberia. A second comparison model was made from all 1980 observatory annual means to test for bias in the palaeomagnetic site distribution; the resulting core field was very similar to that shown in Fig. 2b. Finally we repeated the calculations using all individual data, equally weighted, rather than site averages. The results were essentially the same, so we decided to use averages to match the marine sediment data, which are also averages.

Next we used our model to predict inclinations from marine sediments. No systematic differences were found between predictions and observations and we therefore combined data sets: the resulting core field retains all its main features (Fig. 3a). Next

FIG. 2 a, Radial component of magnetic field at the core surface for a model that fits the normal-polarity lava data (39 inclinations, 39 declinations) to $\sim 1\sigma$. b, Core field from synthetic data using the same site distribution, data accuracy, and model norm as a. The similarity between the two maps in the Northern Hemisphere is striking. (Contour interval in a and b, 10^5 nT, where the field has been normalized to the same dipole moment as 1980.)



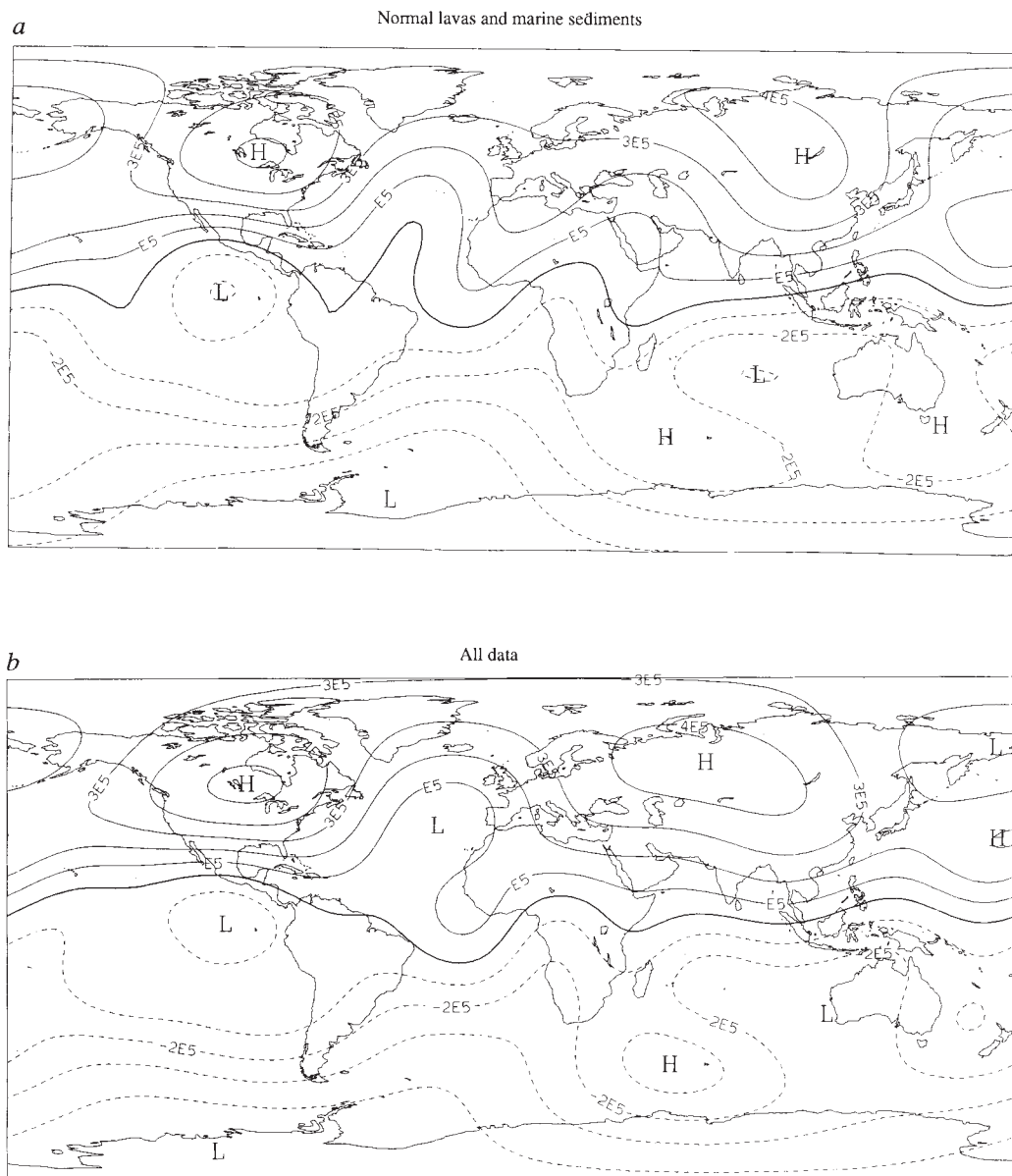


FIG. 3 *a*, Core field from all normal-polarity data (215 inclinations, 39 declinations) produced from both lavas and marine sediments. *b*, Core field obtained by fitting all data (354 inclinations, 53 declinations) to approximately 1σ . Reversed directions have been changed in sign. Spherical harmonic coefficients for this model (to degree 10) are available from the authors on request. (Contour interval as Fig. 2.)

we analysed reversed-polarity data; the core field was found to be smooth and featureless as was a comparison 1980 model, showing there are too few reversed data to determine a satisfactory core field. We then combined both normal- and reversed-polarity data into a single model (Fig. 3*b*); it has a similar appearance to the previous models but more degrees of freedom (33) and smaller formal covariances, reflecting the larger data base.

Modern and palaeofields have significant differences that are not obvious at the core surface. The mean surface energy ($\int \mathbf{B}^2 dS / \int dS$, where the integrals are taken over the Earth's surface) of the Earth's magnetic field, $\mathbf{B}$, minus its axial dipole is 10.0% for the modern field (10.2% for the comparison field in Fig. 2*b*) but only 0.6% for the palaeofield. Time averaging reduces energy in the equatorial dipole by a factor of 80, and reduces energy in degrees two and three by a factor of ten. Degree four terms are not reduced, which is not surprising in view of the 4-fold symmetry in the main lobes. Surface inclination anomalies and declinations are much reduced from modern values because of the time averaging (Fig. 4). The inclination maps support the earlier conclusion³ that 'near-sidedness' (posi-

tive inclination anomaly), rather than the usual far-sidedness, should be searched for in Arctic Canada, Siberia and the Southeast Pacific (but not the South Atlantic).

Three important inferences may be drawn from these results. First, the striking resemblance of the main Northern Hemisphere features in the palaeomagnetic and modern fields is further evidence (see also refs 8, 9) of long-term core-mantle interaction: the morphology of the geomagnetic field is unlikely to persist for more than a few thousand years, the overturn time of core fluid, without the influence of the solid mantle. Similarities in the Southern Hemisphere are much less clear, and the main lobes are virtually absent in the palaeofield (Fig. 3*b*). We suggest that the modern intense secular variation¹ has persisted in this region, perhaps intermittently, throughout the past 2.5 Myr to produce a nearly axisymmetric pattern. The idea of stationary magnetic features tied to the mantle is supported by an idealized numerical model¹⁰ which exhibits a rich variety of dynamic behaviour, including locking, resonance and secondary resonance. It is likely that the Earth's core will respond to mantle forcing in a similarly complex way.

Second, time averaging does not produce axial symmetry. Pre-

Inclination anomaly

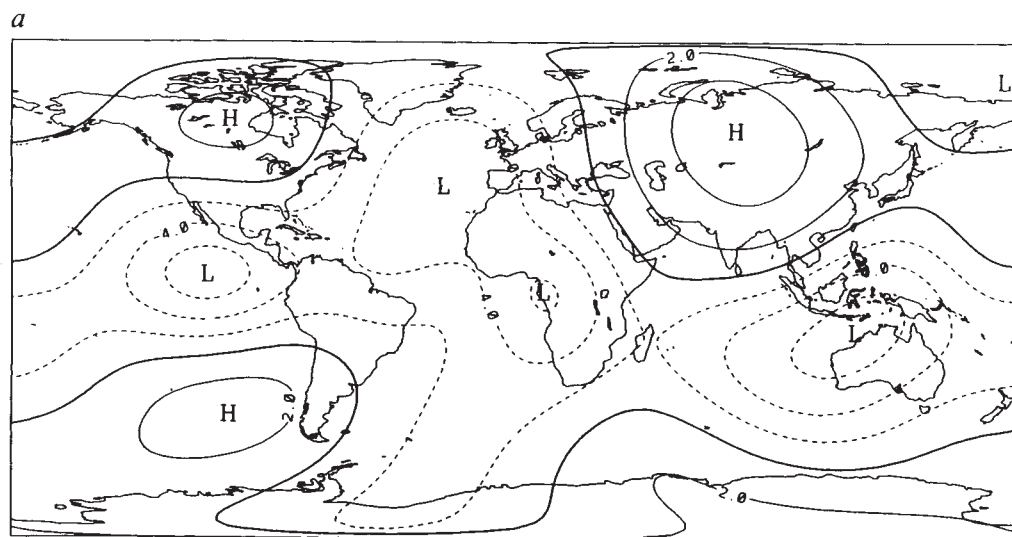
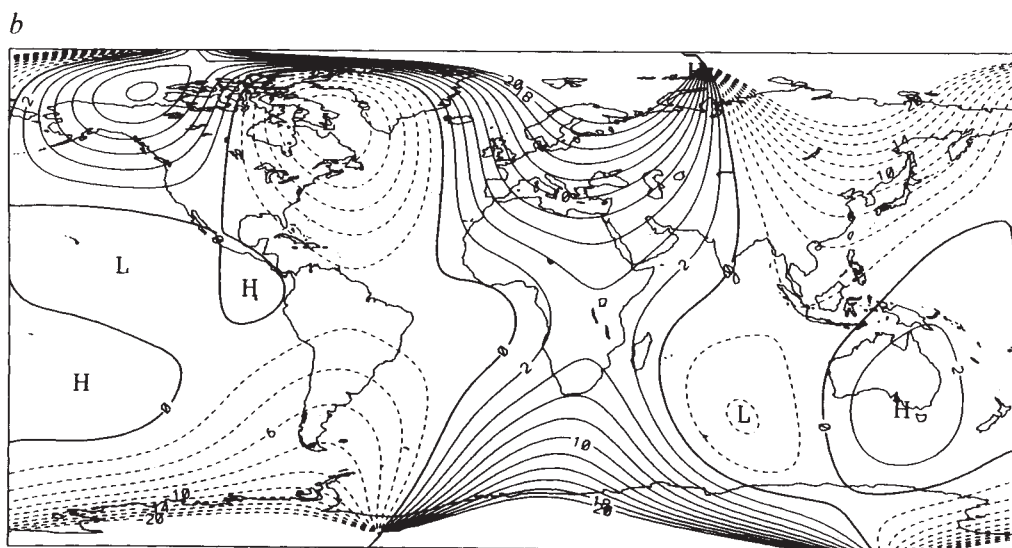


FIG. 4 Inclination anomaly (inclination minus that for an axial dipole) (a) and declination (b) for the final field model at the Earth's surface. (Contour interval in a and b, 2° .) The corresponding core field is shown in Fig. 3b.

Declination



vious studies^{11,12} fitted a few low-order, usually axially symmetric, spherical harmonics to the palaeodata and claimed the residuals lacked statistical significance. The modern field, however, is not restricted to a few low-order harmonics and there is no reason to suppose the palaeofield is any different, particularly as we now know westward drift to be confined in longitude. Departures from axial symmetry also manifest themselves as longitude biases in virtual geomagnetic pole (VGP) positions¹³. A central result of palaeomagnetism is that palaeomagnetic poles, the average of many determinations, lie closer to the geographic pole than the individual VGPs⁷. The present findings show that this arises because time-averaging reduces the equatorial dipole, not because westward drift produces axial symmetry.

Finally, if the concentration of flux towards certain regions of the core-mantle boundary persists during reversal transitions, these concentrations can dictate the longitude of a VGP¹⁴. Our new findings are therefore consistent with the plethora of reversal

transition paths that have recently been found to pass through the Americas or near the antipodal path through Asia^{8,9,15}. □

Received 29 March; accepted 26 August 1993.

1. Bloxham, J. & Gubbins, D. *Nature* **317**, 777–781 (1985).
2. Merrill, R. T. & McElhinny, M. W. in *The Earth's Magnetic Field (Its History, Origin and Planetary Perspective)* (Academic, San Diego, 1983).
3. Gubbins, D. *J. geophys. Res.* **93**, 3413–3420 (1987).
4. Bloxham, J., Gubbins, D. & Jackson, A. *Phil. Trans. R. Soc.* **329**, 415–502 (1989).
5. McFadden, P. L., Merrill, R. T., McElhinny, M. W. & Lee, S. *J. geophys. Res.* **96**, 3923–3933 (1991).
6. Schneider, D. & Kent, D. V. *Ref. Geophys.* **28**, 71–96 (1990).
7. McElhinny, M. W. in *Paleomagnetism and Plate Tectonics* (Cambridge Univ. Press, 1973).
8. Laj, C., Mazaud, A., Weeks, R., Fuller, M. & Herrero-Bervera, E. *Nature* **351**, 447 (1991).
9. Clement, B. M. *Earth planet. Sci. Lett.* **104**, 48–58 (1991).
10. Zhang, K. & Gubbins, D. *J. Fluid Mech.* **250**, 209–232 (1993).
11. Livermore, R. A., Vine, F. J. & Smith, A. G. *Geophys. J. R. astr. Soc.* **73**, 153–171 (1983).
12. Coupland, D. H. & van der Voo, R. *J. geophys. Res.* **85**, 3529–3548 (1980).
13. Constable, C. G. *Nature* **358**, 230–233 (1992).
14. Gubbins, D. & Coe, R. *Nature* **362**, 51–53 (1993).
15. Tric, E. et al. *Phys. Earth planet. Inter.* **65**, 319–336 (1991).

Oldest *Homo* and Pliocene biogeography of the Malawi Rift

Friedemann Schrenk*, Timothy G. Bromage†, Christian G. Betzler‡, Uwe Ring§ & Yusuf M. Juwayeyi||

* Department of Palaeontology, Hessisches Landesmuseum, Friedensplatz 1, 64283 Darmstadt, Germany

† Department of Anthropology, Hunter College, CUNY, 695 Park Avenue, New York, New York 10021, USA

‡ Geologisch-Paläontologisches Institut, Universität Frankfurt, Senckenberg-Anlage 32, 60325 Frankfurt, Germany

§ Institut für Geowissenschaften, Universität Mainz, 55099 Mainz, Germany

|| Department of Antiquities, PO Box 264, Lilongwe, Malawi

THE Malawi Rift and Pliocene palaeofaunas, which include a hominid mandible attributed to *Homo rudolfensis*, provide a biogeographical link between the better known Plio-Pleistocene faunal records of East and Southern Africa. The Malawi Rift is in a latitudinal position suitable for recording any hominid and faunal dispersion towards the Equator that was brought on by increased aridity of the Late Pliocene African landscape. The evidence suggests that Pliocene hominids originated in the eastern African tropical domain and dispersed to southern Africa only during more favourable ecological circumstances.

Formation of the Malawi Rift began about 8 Myr ago. Subsequent rifting and faulting led to subsidence which created a riverine system and eventually a rift lake (palaeolake Malawi) 5–4 Myr ago (Fig. 1)¹. The Plio-Pleistocene Chiwondo Beds include fluvial, palaeosol, swamp, beach, and foreshore and offshore lacustrine deposits (Fig. 2)², and so provide the southernmost African Rift Valley occurrences.

After the pioneering surveys of Malawi Rift faunas^{3–6}, the Hominid Corridor Research Project (HCRP) began a long-term study in 1983 which focused on the role of southeastern Africa in the origin and dispersion of Plio-Pleistocene faunas and early hominids.

The HCRP has identified 131 fossil localities in the Karonga and Uraha areas (Fig. 1), representing two biochronological zones (Fig. 2). The faunal assemblages indicate an age of 4 Myr and older for unit 2 and of 3–1.5 Myr for unit 3 (see Fig. 2 legend).

An early hominid mandibular corpus in two joining parts (UR 501; Fig. 3) was recovered from stratigraphic unit 3A (Fig. 2). The two halves are broken behind the roots of the rami posterior to the RM₂ and through the LM₂. The superior lateral tori fade quickly off the roots of the anterior borders of the rami, and the inferior marginal tori are only weakly developed. A broad jugum incorporates the C and P₃ roots. Mental tubercles and a mental trigone are distinct. An inferior transverse torus is weakly developed at the level of mesial P₄. The mental foramen is situated about 4 mm below corpus midline. P₃s are intact lingually, oval in shape, and with a small distolingually placed talonid. P₄s are well preserved, with a lingual cusp distal to the larger buccal cusp; the large talonid projects distolingually. The five cusps of

FIG. 1 Lithology and structure of the northern Malawi Rift. Right: Tectonic studies of the western^{19,20} and eastern²¹ branches of the East African Rift Valley system show that initial rifting is older in the eastern branch. Malawi rifting (western branch) began ~8.0 Myr ago. An early phase of normal faulting was followed by a strike-slip dominated system. The western side of the Karonga-Chilumba basin was susceptible to localized uplift and thus Cenozoic Lake sediments were exposed. Left: Lithology and structural elements of the northern Malawi Rift, showing border fault segments (arrows) adjacent to lake depocentres (dashed line represents 550 m depth); opposite border fault segments are en-echelon step faults and flexures. Plio-Pleistocene deposits ('Lake Beds' indicates Chiwondo Beds, Chitimwe Beds) occur south of Karonga and north of Chilumba (Uraha).

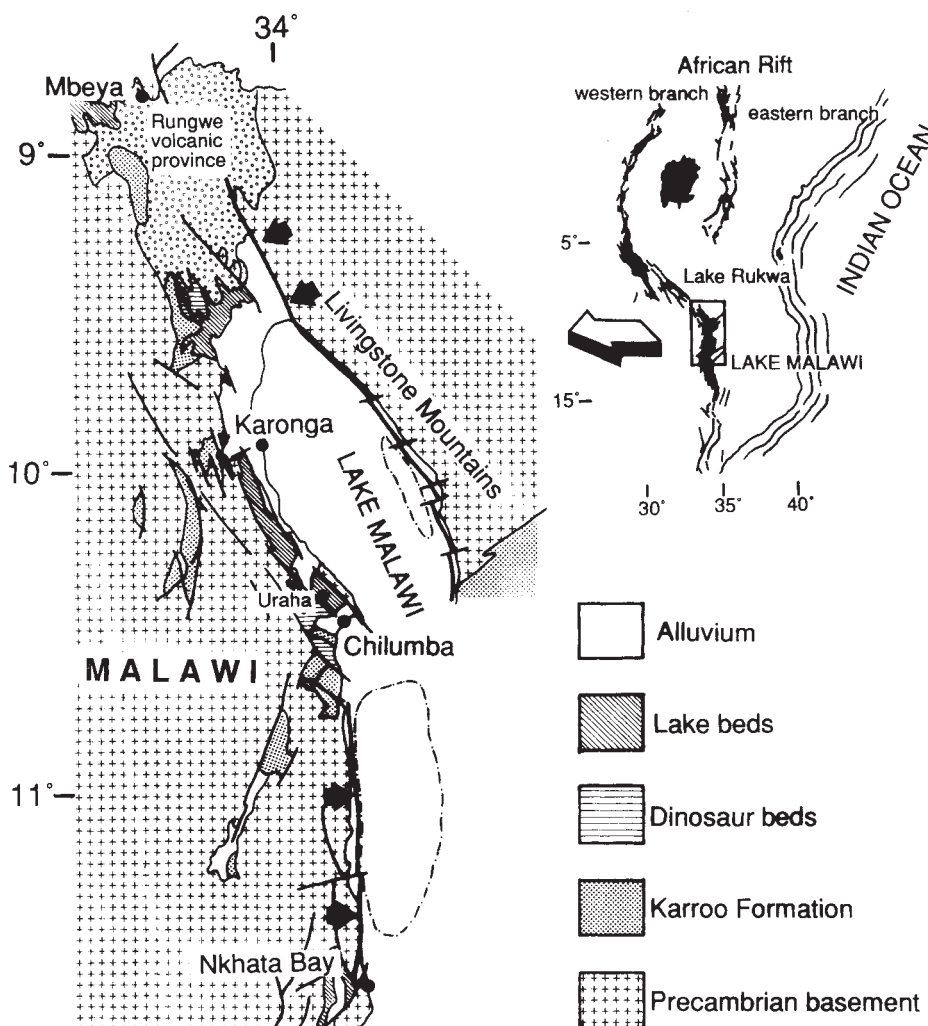


TABLE 1 Mammalian taxa recovered from unit 3A of the Chiwondo Beds, northern Malawi, grouped according to geographic distribution

Southern Africa only	
Artiodactyla	
<i>Notochoerus capensis</i>	
<i>Potamochoeroides shawi</i>	
<i>Gazella</i> sp. aff. <i>vanhoepeni</i>	
<i>Tragelaphus</i> aff. <i>angasi</i>	
Shared southern and eastern Africa	
Primates	
<i>Theropithecus</i> sp.	
<i>Parapapio</i> sp.	
Proboscidea	
<i>Elephas recki</i>	
<i>Mammuthus subplanifrons</i>	
Perissodactyla	
<i>Equus</i> sp.	
<i>Hipparion</i> sp.	
<i>Ceratotherium</i> aff. <i>simum</i>	
<i>Diceros bicornis</i>	
Artiodactyla	
<i>Metridiochoerus andrewsi</i>	
<i>Phacochoerus</i> sp.	
<i>Hippopotamus</i> sp.	
<i>Tragelaphus</i> aff. <i>strepsiceros</i>	
<i>Megalotragus kottwinkeli</i>	
<i>Syncerus</i> sp.	
<i>Connochaetes</i> aff. <i>taurinus</i>	
<i>Giraffa camelopardalis</i>	
<i>Camelus</i> sp.	
Eastern Africa only	
Primates	
<i>Homo rudolfensis</i>	
Proboscidea	
<i>Elephas ekorensis</i>	
<i>Loxodonta adaurora</i>	
<i>Deinotherium bozasi</i>	
Artiodactyla	
<i>Notochoerus euilus</i>	
<i>Kolpochoerus limnetes</i>	
<i>Metridiochoerus compactus</i>	
<i>Ugandax</i> sp.	
<i>Kobus sigmoidalis</i>	
<i>Kobus</i> aff. <i>patulicornis</i>	
<i>Oryx</i> aff. <i>gazella</i>	
<i>Damaliscus</i> sp.	
<i>Aepyceros</i> sp.	
<i>Madoqua</i> sp.	
<i>Giraffa stillei</i>	
<i>Giraffa pygmaea</i>	
<i>Camelus</i> sp.	

FIG. 2 Sedimentary succession in the southern part of the Karonga-Chilumba basin at Uraha. The Plio-Pleistocene Chiwondo Beds²², overlying the Jurassic-Cretaceous dinosaur beds unconformably and are partially covered and eroded by late Pleistocene alluvial fan deposits (Chitimwe Beds). The large-scale transgressive-regressive cycle represents a highly dynamic depositional system in a nearshore to backshore position. Facies elements are fluvial, palaeosol, swamp, beach, and offshore lacustrine deposits. Maximum thickness is 125 m and five depositional sequences (units 1–5) are limited by unconformities (palaeosols, angular unconformities) reflecting lake level changes or tectonic activity. An endemic gastropod species *Bellamya* cf. *pagodiformis* in upper unit 2 abets the relative ordering and correlating of sedimentary units and fossil localities that are distant from one another²³. Age estimates rely on correlation with radiometrically dated biochronological units in eastern Africa^{13,24}. Unit 1 consists of reddish-to-greyish braided stream deposits. No vertebrate fossils occur and no age attribution is possible. Unit 2 reflects the flooding of the depositional area; wave-accumulated bioclastic beaches show an aggrading to strongly prograding stacking pattern. Distally, silts and sands with abundant pelecypods were deposited. Marls with oncoids around gastropod shells characterized the low-energy depositional ramps. The vertebrate fauna indicates an Early Pliocene age; a primitive proboscidean *Anancus kenyensis* and the suid *Nyanzachoerus jaegeri* indicate an age of about 4.0 Myr or older²⁵. Of the Hippotragini, there are specimens similar to those from the Laetoli Beds at 3.6–3.8 Myr⁵. Unit 3, marked at the base by an angular unconformity in the Karonga area and a major palaeosol horizon in the Chilumba area (Uraha Hill), contains two subunits: Unit 3A, meandering river and deltaic deposits, is strongly condensed at Uraha Hill, as reflected by a major ferruginous calcimorphic palaeosol. The hominid lower jaw UR 501 was found within this condensed section. Faunas associated with UR 501 are¹³: *Notochoerus euilus* (3.35–2.5 Myr), *Tragelaphus* sp. aff. *angasi* (3.0 Myr), an early *Notochoerus scotti* (3.0–2.3 Myr), a mid-late Pliocene *Hipparion* (2.9 Myr or younger), *Ceratotherium simum* (2.5 Myr or younger), *Oryx* (3.0–1.64 Myr) and a late Pliocene representative of *Metridiochoerus andrewsi* (2.3–1.9 Myr). This assemblage is similar to that reported from the Chemeron Formation⁹ in association with purported early *Homo* (KNM-BC1), except that *Notochoerus euilus* of unit 3A points to an older age, whereas KNM-BC1 is associated with a younger suid (*N. scotti*) not represented in unit 3A. Although the associated faunas represent a time-transgressive deposit, either an older (~2.5 Myr) or, more conservatively, an age closer to 2.3 Myr is suggested. Thus 2.4 Myr for UR 501 remains a consensus date until further studies can determine an absolute age. Unit 3B, restricted to Uraha Hill, consists of a series of stacked ferruginous calcimorphic palaeosol horizons. A late Pliocene to early Pleistocene date (2 to 1.5 Myr) is indicated by *Tragelaphus* aff. *strepsiceros*, *Connochaetes* aff. *taurinus* and *Metridiochoerus compactus*²⁶. Unit 4, restricted to the southernmost part of the Karonga-

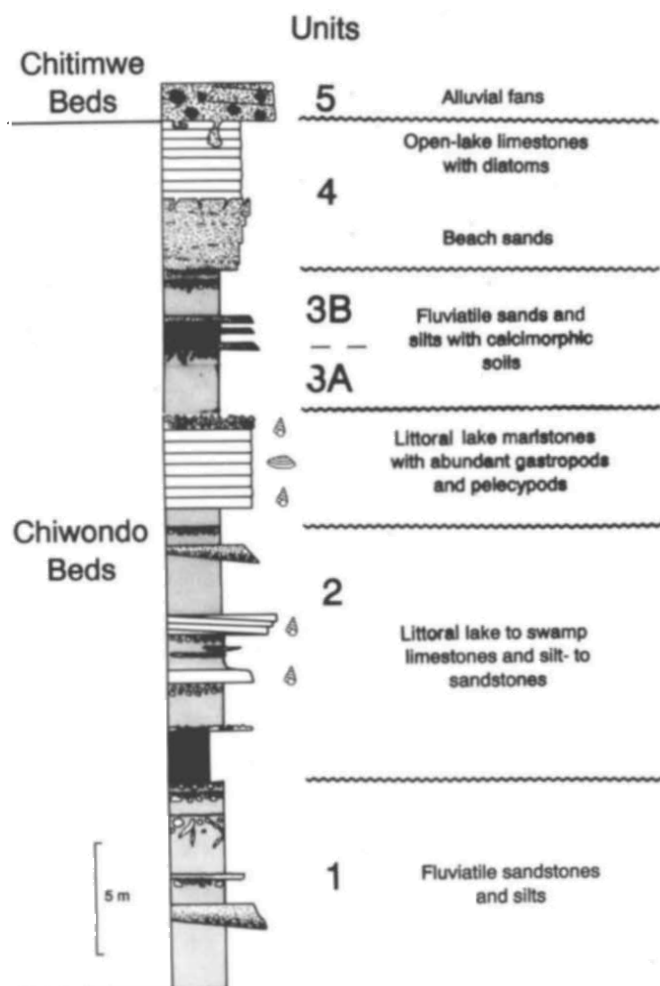
from Koobi Fora, Kenya, such as KNM-ER 1802, in overall corpus dimensions, strength and the morphology of the 'chin' region. These specimens also share absolutely large molar crown areas together with diminished M₂ protoconid relative cusp area, P₃ molarization (that is, relative talonid expansion), plate-like P₃ and P₄ roots, and some enamel microanatomical features that correspond more closely to the *Paranthropus* condition⁷. Taken together, UR 501 corresponds closely to the subset of Late Pliocene fossils from Koobi Fora with relatively large brains and robust jaws and teeth that have been referred to *Homo rudolfensis*⁸, and to which we also refer UR 501.

UR 501 is biochronologically aged at 2.4 Myr (see Fig. 2 legend for unit 3A). Hill *et al.*⁹ have reported the occurrence of an early *Homo* temporal bone (KNM-BC1) traced to radiometrically dated 2.4 Myr deposits from the Chemeron Formation, Kenya. Wood¹⁰ suggests that this specimen may be an early representative of *Homo rudolfensis*, whereas other specimens assigned to this taxon⁸ derive from just below and above the KBS Tuff of the Koobi Fora Formation and date to approximately 1.9–1.8 Myr. Thus a date of 2.4–1.8 Myr is at present indicated for *Homo rudolfensis*.

The significance of UR 501 lies in the context of its temporal

the M₁s are separated by Y-shaped fissure patterns and are worn flat. A low distal marginal ridge surrounds a relatively large talonid. The RM₂ is nearly complete, but only a sheared dentine surface of the left mesial M₂ is preserved. Fissuring demarcates a tuberculum intermedium (C7) platform. The distal longitudinal fissure meets a posterior fovea delimited by a marginal ridge adorned with at least two tubercles of the tuberculum sextum (C6). A distal wear facet indicates that the RM₃ was in occlusion, suggesting that the individual was a young adult.

Many absolute and relative enamel and premolar crown shape indices, relative cusp areas, enamel microanatomical features, fissure patterns and crown morphology are within the range of early *Homo*, though some may be as well subsumed within the limits of variation represented by *Australopithecus* (*A. afarensis* and *A. africanus*)⁷. However, UR 501 is similar to specimens



Chilumba Basin (Uraha), consists of a transgressive part with well-sorted, unconsolidated beach sands overlain by open lake limestones with abundant diatoms. No date has been proposed. Unit 5 (Chitimwe Beds) overlies the Chiwondo Beds (units 1–4) with an erosional unconformity. The formation of the alluvial fans is related to a major lowering of lake level after deposition of unit 4. No reliable age has been proposed.



FIG. 3 Early hominid mandible UR 501 (*Homo rudolfensis*) from Uraha in occlusal view. Scale bar, 1 cm. Selected dental measurements (corrected) of UR 501 are compared with KNM-ER 1802 (ref. 27):

	UR 501		KNM-ER 1802	
	Buccal lingual	Mesial distal	Buccal lingual	Mesial distal
LP ₄	11.4 mm	9.8 mm	11.8 mm	12.1 mm
LM ₁	13.2 mm	14 mm	13.2 mm	14.9 mm
RP ₃	11.4 mm	9.7 mm	11.5 mm	10.7 mm
RP ₄	11.5 mm	9.5 mm	12.0 mm	11.4 mm
RM ₁	13.3 mm	14 mm	13.0 mm	14.8 mm
RM ₂	14.6 mm	17.3 mm	14.2 mm	16.6 mm

Selected corpus measurements of UR 501 are compared with KNM-ER 1802 (ref. 27):

	UR 501	KNM-ER 1802
Symphyseal height	34 mm (estimate)	36 mm
Symphyseal depth (max.)	16.8 mm	24.5 mm
Corpus height at RM ₁	34 mm	38 mm
Corpus width at RM ₁	21 mm	23 mm

origin and its biogeographical relationships. These are demonstrated by placing the mammalian assemblage of unit 3A into three geographically based groups (Table 1). *Homo rudolfensis* is associated with an assemblage dominated by eastern African endemic faunas, whereas the group of southern African endemics of the Malawi Rift is strikingly small. This pattern reflects the equatorward dispersion of southern African faunas, which is in accord with the northward-drifting of vegetation belts during the aridification of global climates at about 2.5 Myr^{11–13}. Evidence for this resides in patterns of extinction and speciation among Pliocene Bovids^{14,15}. Southern African endemics dispersing to eastern Africa contribute to the larger shared 'southern and eastern' faunal group in Table 1. Equatorial lineages faced not so much the problem of latitudinal shifts in biome zonation as habitat fragmentation, and, having a greater diversity in habitat choices, remained endemic to the tropical African ecological domain and experienced either extinction or speciation. We therefore suggest that *Homo rudolfensis* arose during, and partly as a result of, the 2.5 Myr climatic cooling event in eastern Africa and remained endemic there in the face of prevailing equatorward dispersion tendencies in other taxa according to the 'habitat theory'¹⁶.

In southern Africa the supposed time-aggregated deposits of Sterkfontein member 4 (ref. 17) overlap the age of Chiwondo Beds unit 3A but do not represent *Homo rudolfensis*; instead they represent a southern African endemic, *Australopithecus africanus*. As *H. habilis* is present in the Late Pliocene at Koobi Fora, the possibility remains that *A. africanus* dispersed to eastern Africa around 2.5 Myr ago and evolved into *H. habilis* during this phase of equatorward biome constriction. If the assignment is accepted of early *Homo* specimens from Sterkfontein member 5 (ST5) and Swartkrans member 1 (SK1) to *Homo habilis*, then we can attribute this taxon to a shared 'southern + eastern' African group, which according to our ecological model, later expanded its distribution deeper into southern Africa between 1.8–1.5 Myr ago¹⁸ during a period of biome expansion¹³.

The pattern emerging from our biogeographical interpretation is that early hominids persistently arose in the tropical ecological domain of eastern Africa, a pattern that also dominates speciation events of the Bovidae 2.5 Myr ago. Only during more favourable periods did early hominids disperse southwards, evolve and become established in an environmentally temperate ecological domain. □

Received 29 March; accepted 9 August 1993.

- Ring, U. & Betzler, C. *J. hum. Evol.* (in the press).
- Betzler, C. & Ring, U. *J. hum. Evol.* (in the press).
- Clark, J. D., Stephens, E. A. & Coryndon, S. C. *Am. Anthropol.* **68**, 46–87 (1966).
- Clark, J. D., Haynes, C. V. Jr, Mawby, F. E. & Gautier, A. *Quaternaria* **13**, 305–354 (1970).
- Kaufulu, Z. M., Vrba, E. S. & White, T. D. *Ann. Transv. Mus.* **33**, 1–8 (1981).
- Bromage, T. G. & Schrenk, F. *J. hum. Evol.* **15**, 497–500 (1987).
- Bromage, T. G., Schrenk, F. & Zonneveld, F. W. *J. hum. Evol.* (in the press).
- Wood, B. *Nature* **355**, 783–790 (1992).
- Hill, A., Ward, S., Deino, A., Curtis, G. & Drake, R. *Nature* **355**, 719–722 (1992).
- Wood, B. *Nature* **355**, 678–679 (1992).
- Bonnefille, R. in *Earliest man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. I. & Leakey, R. E. F.) 421–431 (Univ. Chicago Press, 1976).
- Vrba, E. S., Burckle, L. H., Denton, G. H. & Partridge, T. C. S. *Afr. J. Sci.* **81**, 224–275 (1985).
- Bromage, T. G., Schrenk, F. & Juwayeyi, Y. M. *J. hum. Evol.* (in the press).
- Vrba, E. S. in *Evolutionary History of the Robust Australopithecines* (ed. Grine, F. E.) 405–426 (de Gruyter, New York, 1988).
- Turner, A. & Wood, B. *J. hum. Evol.* **24**, 147–168 (1993).
- Vrba, E. S. *J. Mamm.* **73**, 1–28 (1992).
- Delson, E. *Cour. Forsch. Inst. Senckenberg* **69**, 199–218 (1984).
- Vrba, E. S. in *Prehyst. 1er Congr. Internat. Paleont. Humaine*, 707–752 (CNRS, Paris, 1982).
- Ebinger, C. *J. Geol. Soc. Am. Bull.* **101**, 885–903 (1989).
- Crossley, R. & Crow, M. *J. Academia Nazionale Lincei* **47**, 77–87 (1980).
- Strecker, M. R., Blisniuk, P. M. & Eisbacher, G. H. *Geology* **18**, 299–302 (1990).
- Dixey, F. Q. *J. geol. Soc. Lond.* **83**, 432–447 (1927).
- Schrenk, F., Gorthner, A. & Bromage, T. G. *J. Hum. Evol.* (in the press).
- Brown, F. H. & Feibel, C. S. in *Evolutionary History of the Robust Australopithecines* (ed. Grine, F. E.) 325–341 (de Gruyter, New York, 1988).
- Brown, F. H. & Nash, W. P. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. L. I. & Leakey, R. E. F.) 50–63 (Univ. Chicago Press, 1976).
- Harris, J. in *Koobi Fora Research Project Vol. 3* (ed. Harris, J. M.) 215–302 (Oxford Univ. Press, 1991).
- Wood, B. *Koobi Fora Research Project, Hominid Cranial Remains Vol. 4* (Oxford Univ. Press, 1991).

ACKNOWLEDGEMENTS. Field work was conducted by permission of the Malawi Government. Funding was provided by Deutsche Forschungsgemeinschaft, L.S.B. Leakey Foundation, National Geographic Society, PSC-CUNY Faculty Award of the City University of New York, Wenner Gren Foundation, Boise Foundation, and Geological Society of America. We thank F. Zonneveld for CT scanning; P. Andrews, R. Bernor, H. B. S. Cooke, E. Delson, M. Köhler, E. Vrba and T. White for discussion and faunal identification; A. Bender and C. Schnubel for drawing Figs 1 and 3 respectively; and M. Leakey (Nairobi), H. Huchinson (Berkeley), B. Rubidge, J. Kitching and P. V. Tobias (Johannesburg) and F. Thackeray (Pretoria) for permission to study faunal and hominid material in their care.

Ultraviolet radiation and coral bleaching

Daniel F. Gleason & Gerard M. Wellington

Program in Evolutionary Biology and Ecology, Department of Biology, University of Houston, Houston, Texas 77204–5513, USA

EPISODES of coral bleaching resulting from dissociation of endosymbiotic algae (zooxanthellae) from host coral tissues have occurred with increasing frequency over the past decade on reefs throughout the tropics^{1,2}. These episodes have usually been attributed to increases in seawater temperatures^{3–10}, but the mass bleaching events that occurred throughout the Caribbean during 1987 and 1990 were not readily explained by temperature alone^{11,12}. An additional factor that may have contributed to these bleaching episodes is ultraviolet radiation in the 280–400-nm band. At many localities where bleaching occurred in 1987 and 1990, sea conditions were described as extremely calm with exceptionally clear water¹³. In the absence of suspended organic and inorganic matter in the water column, higher than average intensities of ultraviolet radiation probably reached all depths within the photic zone for several consecutive months. Evidence for a possible link between ultraviolet radiation and coral bleaching has not been forthcoming². Here we report results of a field experiment showing that, irrespective of high water temperatures, short-term (three weeks) increases in ultraviolet radiation of a magnitude possible under calm, clear water column conditions can readily induce bleaching in reef-building corals.

A large proportion of the Caribbean corals that bleached during 1987 and 1990 were found at depths greater than 20 m. To determine whether increased ultraviolet (UV) radiation in the 280–400-nm band could have contributed to these bleaching

events, we conducted field experiments at San Salvador Island, Bahamas, in September 1991. Ultraviolet light intensities were increased by transplanting the common reef-building species *Montastrea annularis* Ellis and Solander from 24 m to 24, 18 and 12 m depth. At each depth, coral colonies were either exposed to ambient UV radiation or protected from UV by an acrylic cover (Fig. 1). Water temperature was recorded at hourly intervals during the experiment with a DataSonde 3 data logger fitted with a thermocouple. Water temperatures were isothermal between 12 and 24 m, and averaged 29.1 °C (± 0.12 s.d.; range, 28.7–29.4 °C; $n = 665$; continuous hourly measurements).

Transplanted colonies of *M. annularis* exposed to UV radiation at 12 m depth began to show visible signs of bleaching (pigment loss) within seven days of exposure. Twenty-one days after transplantation, all colonies irradiated with UV light at 12 m depth were much more pale than their shielded counterparts. In contrast, changes in colony colour were not observed either for UV-protected colonies at 12, 18 and 24 m or for UV-exposed colonies at 18 and 24 m depth. All coral colonies were collected 21 days after transplantation and were further assayed for the effects of increased UV radiation (Fig. 2a–d). No significant differences in the number of zooxanthellae per cm² and μ g chlorophyll per cm² were detected between treatments at 24 and 18 m depth, but colonies exposed to UV at 12 m gave significantly lower values than 24-m controls (Fig. 2a and b). In contrast, zooxanthella densities and chlorophyll concentrations in colonies shielded from UV radiation at 12 m were not significantly different from the 24-m controls (Fig. 2a and b). These data indicate that responses in UV-exposed corals at 12 m were a result of increased UV rather than of increased visible light (400–700 nm), which photoinhibits free-living phytoplankton^{14,15}. Enhancing UV light intensities did not affect the concentrations of either chlorophyll per zooxanthella cell or sol-

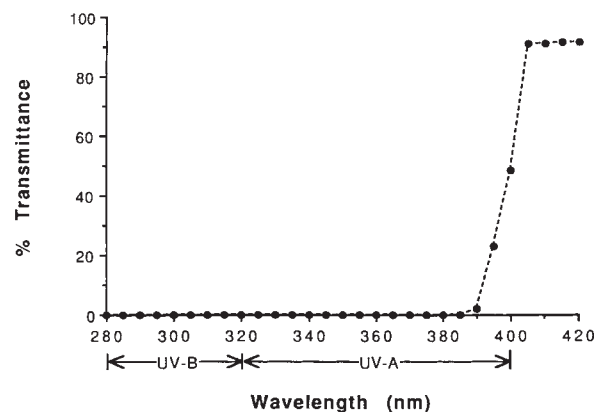


FIG. 1 Ultraviolet light transmission properties of acrylic covers (type OP-2, Cyro Industries) used to shield transplanted colonies of *M. annularis* from UV-A and -B radiation. At wavelengths between 420 and 700 nm, the acrylic has a transmission of 92%. Per cent transmission in the UV region of the spectrum was determined at 5-nm intervals on a Hewlett-Packard model 8452A diode array spectrophotometer. Acrylic covers (30 × 30 × 0.6 cm) were secured to the reef by drilling 6–8 cm deep into the substratum with an air-powered pneumatic drill and cementing 1 × 22 cm stainless steel rods in these holes with underwater epoxy. Each plastic cover was supported at two of its diagonally opposite corners with a stainless steel rod. Acrylic covers were kept free of encrusting organisms and sediment during the experiment to avoid reductions in visible light (400–700 nm) transmission. In addition, we tested for alterations in current velocity which might have resulted from inclusion of the covers by monitoring dissolution of plaster of Paris clods²³. At all three depths, water flow rates beneath acrylic covers were not significantly different from those observed on bare substratum (two-tailed independent samples *t*-test, $P = 0.32$, 0.20 and 0.17 for 12, 18 and 24 m depth, respectively).

TABLE 1 Integrated UV-A and -B light intensities between 11:00 and 14:00 h (DST) at three depths on San Salvador Island, Bahamas

Depth (m)	n	Ultraviolet intensity (W m^{-2})	
		UV-B (300–320 nm)	UV-A (320–400 nm)
12*	17	0.32 (.03)	15.70 (1.08)
18*	16	0.12 (.01)	10.06 (1.01)
24*	18	0.05 (.01)	7.59 (0.73)
24†	11	0.10 (.01)	11.44 (1.25)
24‡		0.19	18.17
24§		0.24	22.16

Ultraviolet light was monitored with four LiCor LI-1800UW scanning spectroradiometers (LiCor, Lincoln, Nebraska) deployed at the surface, and at 12, 18 and 24 m depth. Sensors were programmed to scan from 300 to 700 nm in 2-nm increments once every hour between 07:00 and 17:00 h. Values, except for maxima, represent means ($\pm$ s.e.). n is the number of sampling days.

* Intensities observed during transplant experiments (September–October 1991).

† Intensities observed during time of year when intensities are highest (July–August 1991).

‡ Maximum intensities observed during July–August 1991.

§ Theoretically derived maxima (see text for details).

uble protein per cm^2 of coral tissue surface (Fig. 2c and d).

Reef-building corals neutralize the damaging effects of UV light through synthesis or accumulation of UV-absorbing compounds, known as mycosporine-like amino acids (MAA)^{16–18}. As increased intensities of UV radiation induced bleaching in our experiments, we investigated the possibility that *M. annularis* minimizes the impact of this radiation by rapidly enhancing MAA concentrations. Colonies of *M. annularis* showed a gradual reduction in MAA concentrations with increasing depth, indicating that deeper-water colonies (at >20 m) may be particularly vulnerable to sudden increases in UV radiation (Fig. 3a). Measurement of UV-absorbing pigments from colonies transplanted from 24 m to 24, 18 and 12 m depth also indicated that *M. annularis* is unable to counter rapid increases in UV intensity through enhanced synthesis or accumulation of MAAs over a period of 21 days (Fig. 3b).

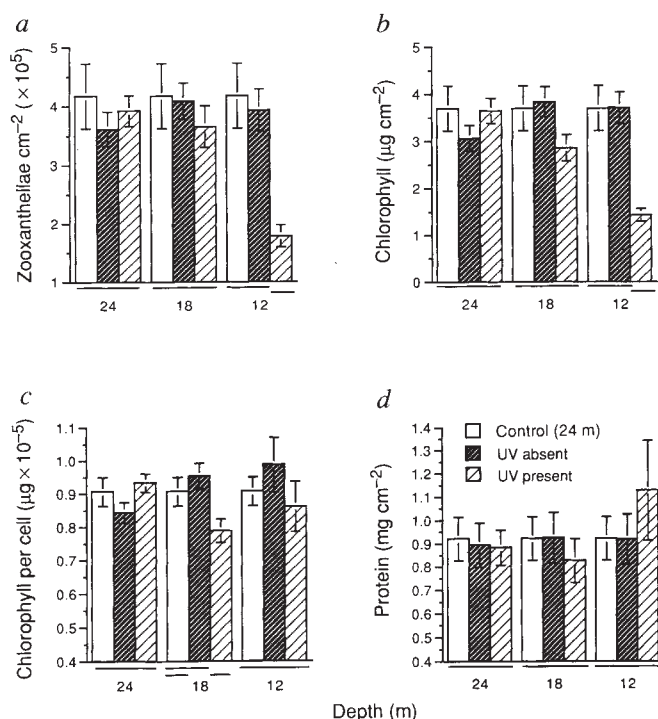
Our transplant experiments, although conducted in September/October, were designed to simulate light intensities that are possible under exceptionally clear water conditions during the time of year when UV radiation is at its peak in northern tropical latitudes (June–August). The extent to which transplantation increased UV radiation levels was estimated by comparing light intensities observed during the experiments with those present in summer. Mean integrated intensities of UV-A (320–400 nm) and UV-B (300–320 nm) at 12 m depth during the experiment were 1.4- and 3.2-fold higher, respectively, than the average irradiances at 24 m during the summer of 1991 (Table 1). The maximum integrated intensities of UV-A and -B (that is, days with exceptionally clear and calm water conditions) measured at 24 m depth during the summer were 116 and 59%, respectively, of those levels responsible for coral bleaching in colonies transplanted to 12 m (Table 1). These empirically derived maxima represent a conservative estimate of the highest UV light levels possible at 24 m. On the basis of the attenuation coefficients for the clearest oceanic waters¹⁹, and the highest UV irradiances observed at 1 m depth during the summer of 1991, we calculated that UV-A and -B intensities at 24 m on reefs in San Salvador can reach 141 and 75% of the mean UV levels responsible for bleaching at 12 m depth (Table 1). Thus, changes in mean UV-A intensity of a magnitude greater than that required to induce bleaching can occur at 24 m depth, and UV-B intensity approaches that which is sufficient to induce bleaching.

It should be noted that our instruments were limited to measuring UV-B intensities at 300–320 nm. Given the small quantities of radiation between 280 and 300 nm reaching the surface of the Earth at northern tropical latitudes²⁰, combined with the high attenuation coefficients for these wavelengths in water²¹, we calculated that this portion of the spectrum contributed less than 1% to our integrated UV-B values even at the shallowest depth (12 m). Therefore, wavelengths between 280 and 300 nm were unlikely to have contributed significantly to the observed bleaching.

We drew three conclusions from our field manipulations and *in situ* UV monitoring. First, increases in UV intensity that result from sudden and prolonged changes in water column clarity

FIG. 2 Effects of UV radiation on (a) zooxanthella densities; (b) chlorophyll concentrations per cm^2 of coral tissue surface area; (c) chlorophyll concentrations per zooxanthella cell; and (d) mg soluble protein per cm^2 coral tissue surface area in *M. annularis* colonies transplanted from 24 m to 24, 18 and 12 m depth. At each depth, colonies were either exposed to (UV present) or protected from (UV absent) ambient UV radiation. Controls were unmanipulated colonies from 24 m. Bars represent means ($\pm$ s.e.). For a, b and c, $n=20$ colonies per treatment, except for controls ($n=10$). For d, $n=10$ colonies in treatments and controls. Treatments connected by underlines were not significantly different at $\alpha=0.05$ (one-way ANOVA and Tukey's HSD test).

METHODS. Zooxanthella densities, chlorophyll concentrations per cm^2 , and chlorophyll concentrations per zooxanthella were quantified on 9- cm^2 samples of coral surface collected from each colony in 100 ml filtered sea water. Zooxanthellae were concentrated by centrifugation and resuspended in 10 ml filtered sea water. Zooxanthellae from three subsamples of this homogenate were counted on a haemocytometer. After determination of zooxanthella densities, the homogenate was recentrifuged. Chlorophyll was extracted from each algal pellet with acetone at 4 °C for 24 h. Pigment concentrations were extrapolated from spectrophotometric measures of chlorophyll-a (663 nm) and chlorophyll-c₂ (630 nm)²⁴. Protein was solubilized by placing tissue sections of coral of 1–2 cm^2 surface area in 2 ml 1M NaOH and heating the sample for 30 min at 90 °C in a water bath. Soluble protein was quantified by spectrophotometry with the dye Coomassie brilliant blue²⁵. Protein concentrations were standardized to tissue surface area as quantified by the aluminium foil technique²⁶.



may contribute substantially to coral bleaching. UV-induced bleaching happens rapidly (within three weeks) and can occur in the absence of raised seawater temperatures. Second, concentrations of MAAs within the coral-algal symbiosis cannot be augmented at a rate sufficient to counter UV light increases resulting from rapid changes in water-column conditions. Third, coral bleaching induced by UV radiation results from reductions in zooxanthella densities rather than the combined effects of zooxanthellae expulsion and decreases in chlorophyll concentrations per algal cell, as has been noted in temperature-related bleaching¹⁰.

Verification that increased UV radiation induces coral bleaching in the absence of abnormally high seawater temperatures raises additional concerns regarding potential climate change

and coral stress. Increases in UV intensity due to anthropogenic ozone depletion will be minimal at low latitudes and should not be a factor in bleaching²², but an increase in the frequency or duration of exceptionally calm (doldrums) and clear water conditions, which accompany such large-scale climatic phenomena as El Niño/Southern Oscillation, for example, could lead to greater penetration of UV radiation on coral reefs and thus to an increased potential for UV-induced bleaching². □

Received 11 February; accepted 23 July 1993.

- Glynn, P. W. *Trends Ecol. Evol.* **6**, 175–179 (1991).
- Glynn, P. W. *Coral Reefs* **12**, 1–17 (1993).
- Jaap, W. C. *Bull. mar. Sci.* **29**, 414–422 (1979).
- Glynn, P. W. *Environ. Conserv.* **11**, 133–146 (1984).
- Atwood, D. K. et al. *Proc. Assoc. mar. Lab. Caribbean* **21**, 47 (1988).
- Glynn, P. W. *A. Rev. Ecol. Syst.* **19**, 309–345 (1988).
- Porter, J. W., Fitt, W., Spero, H., Rogers, C. & White, M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9342–9346 (1989).
- Cook, C. B. et al. *Coral Reefs* **9**, 45–49 (1990).
- Jokiel, P. L. & Coles, S. *Coral Reefs* **8**, 155–162 (1990).
- Glynn, P. W. & D'Croz, L. *Coral Reefs* **8**, 181–191 (1990).
- Goenaga, C., Vicente, V. P. & Armstrong, R. A. *Carib. J. Sci.* **25**, 59–65 (1989).
- Atwood, D. K., Hendee, J. & Mendez, A. *Bull. mar. Sci.* **51**, 118–130 (1992).
- Williams, E. H. & Bunkley-Williams, L. *Atoll. Res. Bull.* **335**, 1–71 (1990).
- Marra, J. *Mar. Biol.* **46**, 203–208 (1978).
- Neale, P. J. in *Photoinhibition* (eds Kyle, D., Osmond, C. & Arntzen, C.) 35–65 (Elsevier, Amsterdam, 1987).
- Jokiel, P. L. & York, R. H. *Bull. Mar. Sci.* **32**, 301–315 (1982).
- Dunlap, W. C. & Chalker, B. E. *Coral Reefs* **5**, 155–159 (1986).
- Gleason, D. F. *Limnol. Oceanogr.* (in the press).
- Smith, R. C. & Baker, K. *Appl. Optics* **20**, 177–184 (1981).
- Gibson, J. H. *Justification and Criteria for the Monitoring of Ultraviolet Radiation* (National Atmospheric Deposition Program/National Trends Network, Colorado State Univ., Fort Collins, 1991).
- Smith, R. C. & Baker, K. S. *Photochem. Photobiol.* **29**, 311–323 (1979).
- Stolarski, R. R. et al. *Science* **256**, 342–349 (1992).
- Wellington, G. M. *Oecologia* **52**, 311–320 (1982).
- Jeffrey, S. W. & Humphrey, G. *Biochem. physiol. Pfl.* **167**, 191–194 (1975).
- Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
- Marsh, J. A. *Ecology* **51**, 255–263 (1970).

ACKNOWLEDGEMENTS. We thank D. Gerace and the Bahamian Field Station for logistical support, and L. Gutierrez, M. Hill and especially J. Schmerfeld for help in the field and laboratory. Supported by the NSF and the National Oceanic and Atmospheric Administration's National Undersea Research Center at the University of North Carolina at Wilmington.

Reverse transduction measured in the isolated cochlea by laser Michelson interferometry

Fabio Mammano & Jonathan F. Ashmore*

Department of Physiology, School of Medical Sciences, University Walk, Bristol BS8 1TD, UK

It is thought that the sensitivity of mammalian hearing depends on amplification of the incoming sound within the cochlea by a select population of sensory cells, the outer hair cells. It has been suggested that these cells sense displacements and feedback forces which enhance the basilar membrane motion by reducing the inherent damping of the cochlear partition^{1–7}. In support of this hypothesis, outer hair cells show membrane-potential-induced length changes^{1–3} at acoustic rates. This process has been termed 'reverse transduction'. For amplification, the forces should be large enough to move the basilar membrane. Using a displacement-sensitive interferometer⁸, we tested this hypothesis in an isolated cochlea while stimulating the outer hair cells with current passed across the partition. We show here that the cochlear partition distorts under the action of electrically driven hair cell length changes and produces place-specific vibration of the basilar membrane of a magnitude comparable to that observed near auditory threshold (about 1 nm). Such measurements supply direct evidence that cochlear amplification arises from the properties of the outer hair cell population.

* To whom correspondence should be addressed.

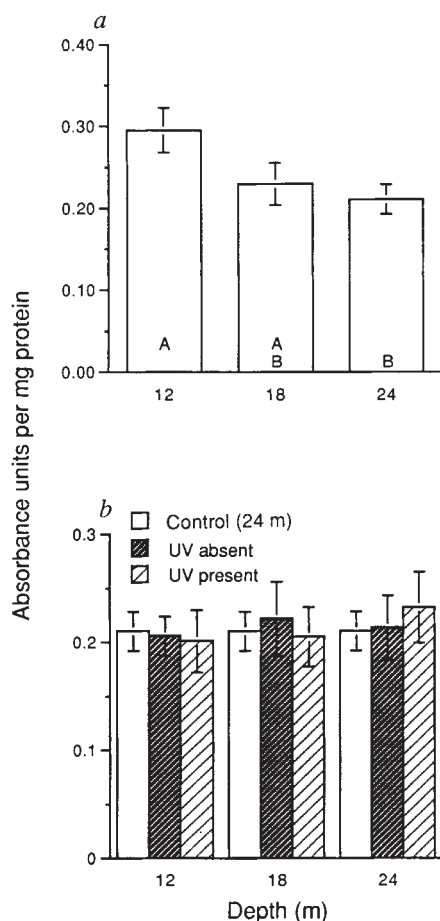


FIG. 3 Absorbance at 320 nm by mycosporine-like amino acids (MAAs) extracted from colonies of *M. annularis*. a, Absorbance units per mg soluble coral protein for unmanipulated colonies collected from 12, 18 and 24 m depth. Values represent means ($\pm$ s.e.) of 10 colonies at each depth. Bars labelled on the interior with the same letter (A or B) were not significantly different at $P=0.05$ (Tukey's HSD test). b, Absorbance units per mg soluble coral protein for colonies transplanted from 24 m to 24, 18 and 12 m depth. Treatments are described in Fig. 2 legend. Bars represent means ($\pm$ s.e.) of 10 colonies. No significant differences in absorbance were detected between treatments at each depth (one-way ANOVA, $P>0.05$ in each case). In both a and b, MAAs were extracted from 1–2-cm² sections of coral tissue with three 5-ml volumes of 20% tetrahydrofuran in methanol. The initial extraction was for 24 h at 4 °C, and the other two were for 5 min under sonication in an ice bath. To detect MAAs, solutions were scanned between 280 and 400 nm on a Milton Roy Spectronic 3000 diode array spectrophotometer. *M. annularis* colonies all had a peak absorbance at 320 nm. Soluble protein was determined after extraction of MAAs as for Fig. 2.

Under visual control, focal electrical stimuli were applied across the organ of Corti by creating a diverging field between a 5 μm -diameter stimulus pipette placed in scala media and an indifferent earth below the isolated cochlea. To obtain access to the third turn of the cochlear spiral, about 11 mm from the stapes, the cochlea was opened and Reissner's membrane gently peeled away. Nanometre displacements of the intact cochlear partition were detected at several positions on the partition surface facing scala media using a displacement-sensitive laser interferometer (Fig. 1a). Placing the pipette tip near the outer hair cells maximized the size of the responses (which were generally too small to be measured if the pipette tip was displaced radially by more than 500 μm from the outermost row of cells).

Figures 1b and 2 show responses of the basilar membrane and those of the reticular lamina (the top surface of the organ of Corti) near the outer hair cells at the same site in cochlear turn 3. The difference in current flowing into the pipette through the outer hair cells and through extracellular space resulted in a net cell depolarization (Fig. 2a) and produced basilar membrane step displacements towards the scala media (Fig. 2b), whereas the reticular lamina moved in the opposite direction (Fig. 2c). The basilar membrane displacements were accompanied by damped oscillations. The motion of the reticular lamina was 5–10 times larger (Figs 1b and 2d) and generally showed an over-damped response (Figs 1b and 2c). The polarity of these responses was reversed by reversing the stimulus current. The steady displacement of the structures indicate that the forces generated by electrical stimuli were large enough to overcome the partition stiffness and to distort its geometry (Fig. 1c). There was no significant difference in basilar membrane responses when the experiments were conducted at 37 $^{\circ}\text{C}$ (data not shown).

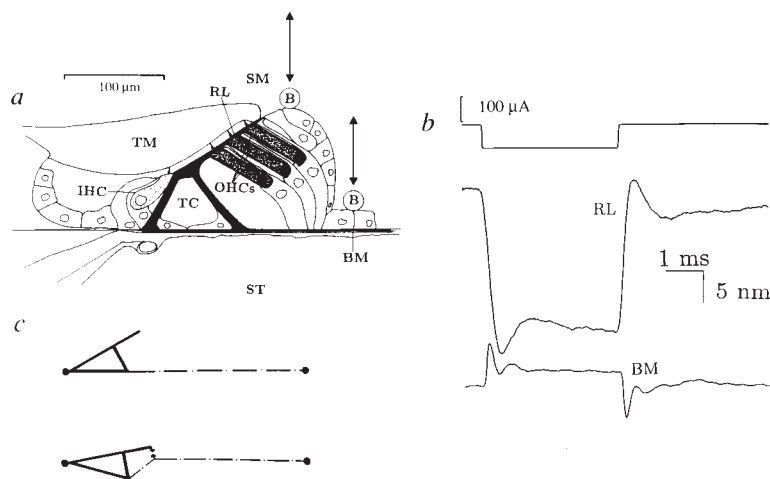
Although no cochlear cells other than the outer hair cells have

been reported to respond mechanically to electrical stimuli, we attempted to single out the origin of the forces elicited by extrinsic current by pharmacological techniques. It is known that salicylate (aspirin) can induce a reversible 20–27 decibel rise in the threshold of response for cochlear nerve fibres at their characteristic frequency^{12,13} and suppress spontaneous otoacoustic emissions¹⁴. Superfusion with 10 mM sodium salicylate reversibly diminishes electromotility in isolated outer hair cells¹⁵. Figure 3 shows that, on introducing an iso-osmotic medium containing 10 mM sodium salicylate for 5 min, a 50% reduction in the size of the basilar membrane response was observed. The response recovered within 12 min. Gadolinium, a reversible blocker of outer hair cell motility¹⁶, was not effective at 1 mM, possibly because of its limited diffusion into the organ of Corti.

These results have implications for cochlear mechanics. The microelectrode recordings show that outer hair cells depolarize, and hence shorten³, in response to negative current pulses. As outer hair cells are strategically inserted between the basilar membrane and the reticular lamina, a change in their resting length would alter their axial tension. Therefore the simplest explanation of our findings is that the organ of Corti alters its dimensions when the outer hair cells change length. As the mean sensitivity of outer hair cells to extrinsic current was -24.8 V/A ($n=5$, range -15.3 to -43.2 V/A), 100 μA would have altered the membrane potential by 2.5 mV, which at the measured membrane potential would cause isolated outer hair cells to change length by 50 nm^{3,17}. The corresponding differential displacement of the cochlear surface by 20 nm implies that the assembly of outer hair cells has a mechanical impedance which numerically matches that of the rest of the cochlear partition. We thus conclude that outer-hair-cell length changes may be functionally less constrained by the surrounding cells than cochlear structure would suggest.

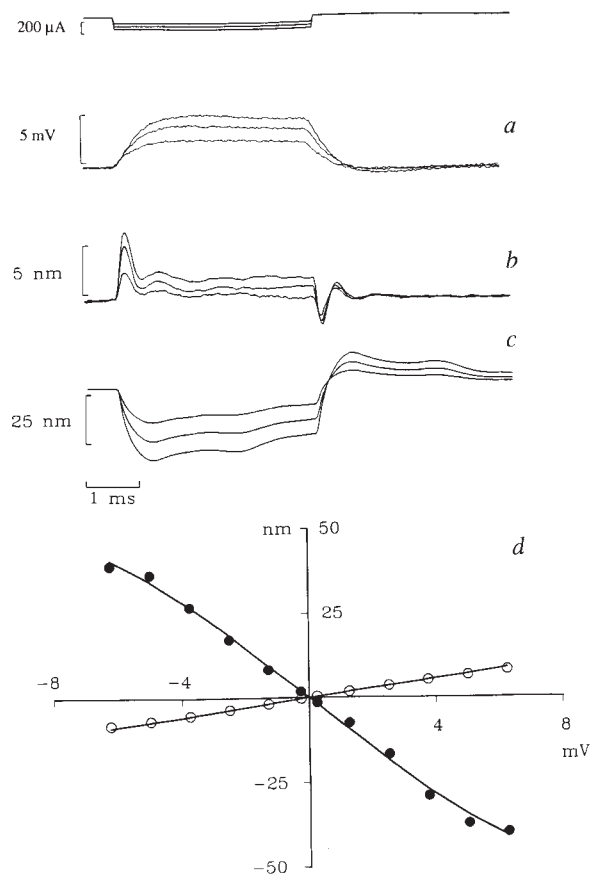
FIG. 1 Schematic drawing of the cochlear partition cross-section, showing the organ of Corti and the position of the measurements. a, Location of the beads (B) where displacement measurements were made is indicated by double vertical arrows. The partition separates scala media (SM) from scala tympani (ST); OHC (IHC), outer (inner) hair cells; TM, tectorial membrane; TC, triangle of Corti, formed by the inner and outer pillar cells. b, Sample responses from the reticular lamina (RL, upper trace) and the basilar membrane (BM, pars pectinata, lower trace) are also shown. For all traces shown in this paper, upward (downward) deflections indicate displacements towards scala media (scala tympani). Note different amplitudes of the motions. After the initial transient following the step onset, both responses settled to a new steady level, but in opposite directions. This implies that the geometry of the partition as well as the basilar membrane tension are altered by internal forces acting within the partition. c, Hypothesized rotation of the TC (heavy lines) around pivot beneath the IHC and induced by OHC shortening.

METHODS. Adult guinea-pigs (200–300 g) were killed by rapid cervical dislocation. After isolating the temporal bone, the bulla was opened and firmly held in a perspex chamber perfused with oxygenated L-15 medium (Gibco) adjusted to pH 7.35 and osmolarity $317 \pm 2 \text{ mosM}$. The cochlear partition was observed under low magnification from scala media by gently scraping away the bony wall of the cochlea and removing Reissner's membrane. Interferometric measurements were made with a $\times 40$ 0.75 NA water-immersion objective (Zeiss). The Michelson interferometer⁸ was illuminated by a stabilized He-Ne 1 mW laser (Model 200, Coherent, Cambridge, UK) and consisted of a fixed reference path and one path in which the beam was reflected back through the objective from aluminium sputter-coated 10 μm glass beads (Mo-Sci Corporation, Rolla, MO, USA) which were allowed to settle on the partition. The intensity of the recombined interfering beams was measured by a photodiode with a flat response from DC to 20 kHz. The diode output was sampled at 31 kHz and averaged 100 times. The stimulus repetition rate was either 8 or 21 Hz. The



interferometer was kept in quadrature under software control. Optimally adjusted, the interferometer was capable of measuring static displacements of the spherical beads. Small radial rocking motions of the beads could cause changes greater than 3 nm in the baseline and in these cases the data were discarded. Displacements were calibrated for each averaging cycle by displacing the reference mirror of the interferometer by a fixed distance, usually 160 nm. Current generated by a controlled stimulator was injected through a micropipette (diameter 5 μm) filled with agar-saline to eliminate fluid bulk flow. Whenever necessary, the stimulator wave-shaping circuitry was used to compensate for the stray capacitance of the pipette so as to reduce the step rise time to less than 20 μs . The current was measured with a virtual ground circuit. Downward current traces correspond to current flowing into the pipette. Temperature, 24–28 $^{\circ}\text{C}$.

FIG. 2 The cochlear partition moves under millivolt changes of hair membrane potential. *a*, Changes in membrane potential induced by current steps in an outer hair cell located underneath the stimulus pipette. Upward deflections indicate cell depolarization. Resting potential, -19 mV. The slowed time course of the intracellular responses is due to the high resistance microelectrode time constant⁹ (1 ms). *b*, Displacements of the basilar membrane elicited by current steps during the same experiment. *c*, Displacement of the reticular lamina in the same cochlea at the same site of *b*. *d*, Voltage-displacement characteristics. Open symbols, basilar membrane; filled symbols, reticular lamina. Abscissae show cell membrane potential changes obtained by converting applied current levels with a sensitivity factor of 25 V/A. Cell depolarization corresponds to motion of the basilar membrane towards scala media and motion of the reticular lamina towards scala tympani. Both responses began to saturate with membrane potential changes in excess of 5 mV (equivalent to 200 μ A stimulus). METHODS. Microelectrodes were pulled from 1 mm (outside diameter) fibre-filled glass tubing (Clark Electromedical, UK) on a P-87 puller (Sutter Instruments) and had a resistance of 150–200 megohms in the bath when filled with 3 M potassium acetate. After completing a series of interferometric measurements, the tectorial membrane was removed with a finely etched tungsten needle held in a manipulator (Singer) while the stimulus pipette was left in place. Outer hair cells in the proximity of the recording site were impaled from the reticular lamina side under direct visual control and the series of stimulus pulses repeated twice for each cell, once with the microelectrode inside and once just above the cell. The microelectrode amplifier output was sampled at 31 kHz and averaged 60 times, (repetition rate 21 Hz). Changes in membrane potential were measured by subtracting the recordings taken inside from those taken outside each cell. Control experiments obtained by passing the electrode through RL showed that the potential difference across RL was negligible (about 0.5 mV at 200 μ A; data not shown). This was a consequence of placing the cochlea in bathing medium which provides a low-resistance pathway to the extracellular current flow. The recorded potential difference between the cell interior and the surface of RL is thus a good estimate of the potential of the basolateral membrane where the OHC motors are located². The average resting membrane potential was -19.2 ± 4.7 mV (s.d., $n=5$). Cells with a resting potential more positive than -14 mV were discarded. Depolarized resting potentials arise when the apical outer hair cell surface



is exposed to artificial perilymph following the removal of Reissner's membrane¹⁰; electromotile responses are, however, insensitive to such altered bathing medium on this surface¹¹.

The antiphase motion of the basilar membrane and the reticular lamina show that each transverse section of the cochlear partition possesses at least two degrees of freedom, so that the top and bottom of the organ of Corti move independently. This represents a departure from classical models of cochlear excita-

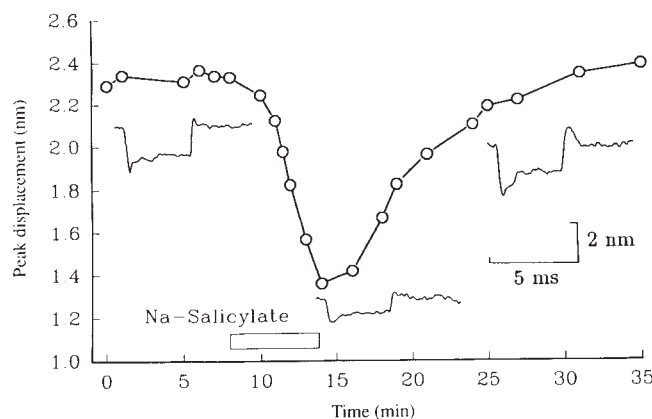


FIG. 3 Partition motion is reduced by salicylate. 10 mM sodium salicylate (bar) in the perfusion medium was added for the time indicated. The response amplitude (and the tuning quality factor) were reduced by 50%.

METHODS. The solution in the 1 ml chamber was continuously changed (2 ml min^{-1}). The estimated dead-time for solution change was about 2 min. To avoid tissue swelling, all solutions were iso-osmotic with control.

tion (ref. 18 for example), which regard the organ of Corti as a single-hinged and rigid structure, but is consistent with those that recognize multiple structures of the partition¹⁹. As proposed in a recent model²⁰, if one mechanical degree of freedom is associated with the basilar membrane pars pectinata and a second with the rigid framework formed by the triangle of Corti and the reticular lamina, the triangle of Corti would rotate towards scala tympani (downwards, Fig. 1) if the outer hair cells shorten and pull the reticular lamina towards the basilar membrane (Fig. 1c). Conversely, the triangle would rotate towards scala media (upwards) if the cells elongate. In both cases, measuring points on the basilar membrane and reticular lamina would move in antiphase when the cells are stimulated.

In such a scheme, the mechanical behaviour of each transverse section can be modelled if we note that the triangle of Corti/basilar membrane system (Fig. 1a, c) can be represented by a pair of masses m_1 , m_2 held by (damped) springs with stiffness k_1 , k_2 , respectively. If these resonant systems are excited simultaneously by equal and opposite forces (applied by the outer hair cells) they would move by amounts inversely proportional to the associated spring stiffness. If the masses are comparable (that is $m_1 \approx m_2$), the ratio of the natural frequency of the oscillators would be proportional to $(k_1/k_2)^{0.5}$. The basilar membrane was tuned to a frequency of 2.3 ± 0.15 kHz (s.d., $n=11$) about 11 mm from the cochlear base, in agreement with published cochlear frequency maps²¹. The responses of the reticular lamina at the same site had a frequency of 1.0 ± 0.15 kHz (s.d., $n=11$). This suggests that $k_1/k_2 = 0.18$ at this point. Thus, the reticular lamina should move about 5.4 times more than the basilar membrane, in reasonable agreement with our data at this cochlear site.

As no sound stimuli were used in these experiments, we are

reporting the effects of outer hair cell forces in the absence of forward mechano-electrical transduction. Cochlear amplification switches in only when these movements are fed back to augment the sensory input to the hair cells^{4,6}. Nevertheless, the cochlear responses reported here show comparable tuning to those induced by sound in a similar preparation²² where a heterodyne interferometer, constructed around an optical sectioning microscope and sensitive to velocity, detected light backscattered from the cochlear partition. Our results support the hypotheses^{4,6,20} that outer hair cells are responsible for amplification within the cochlea as well as for electrically evoked otoacoustic emissions from the intact cochlea²³, and show that it is possible to bridge the gap between *in situ*^{24,25} experiments and *in vivo*²⁶⁻²⁸ observations. □

Received 27 May; accepted 31 August 1993.

1. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ripaubbier, Y. *Science* **227**, 194-196 (1985).
2. Dallos, P., Evans, B. N. & Hallworth, R. *Nature* **350**, 155-157 (1991).
3. Ashmore, J. F. *J. Physiol.* **388**, 323-347 (1987).
4. Neely, S. T. & Kim, D. O. *J. acoust. Soc. Am.* **79**, 1472-1480 (1986).
5. Dallos, P. in *Auditory Function: Neurobiological Bases of Hearing* (eds Edelman, G. M., Gall, W. E. & Cowan, W. M.) 153-188 (Wiley, New York, 1988).

6. Hubbard, A. *Science* **259**, 68-71 (1993).
7. Davis, H. *Hearing Res.* **9**, 79-90 (1983).
8. Mammano, F. & Ashmore, J. F. *J. Physiol.* **452**, 169P (1992).
9. Dallos, P. *Hearing Res.* **14**, 281-291 (1984).
10. Ashmore, J. F. & Meech, R. W. M. *Nature*, **322**, 368-371 (1986).
11. Evans, B. N., Hallworth, R. & Dallos, P. *Soc. Neurosci. Abstr.* **14**, 800 (1988).
12. Evans, E. F., Wilson, J. P. & Borerwe, T. A. 'Tinnitus'. *Ciba Foundation Symposium* **85**, 108-138 (1981).
13. Stypulkowski, P. H. *Hearing Res.* **46**, 113-143 (1990).
14. McFadden, D. & Plattsmier, H. S. *J. acoust. Soc. Am.* **76**, 443-448 (1984).
15. Shehata, W. E., Brownell, W. E. & Dieler, R. *Acta otol.* **111**, 707-718 (1991).
16. Santos-Sacchi, J. J. *Neurosci.* **9**, 2954-2962 (1989).
17. Santos-Sacchi, J. J. *Neurosci.* **11**, 3096-3110 (1991).
18. Davis, H. *Laryngoscope* **68**, 359-382 (1958).
19. Kolston, P. J., de Boer, E., Viergever, M. & Smoorenberg, G. *J. acoust. Soc. Am.* **86**, 133-140 (1989).
20. Mammano, F. & Nobili, R. *J. acoust. Soc. Am.* **93**, 3320-3332 (1993).
21. Greenwood, D. D. *J. acoust. Soc. Am.* **87**, 2592-2605 (1990).
22. Ulfendahl, M., Khanna, S. M. & Flock, A. *Hearing Res.* **40**, 55-64 (1989).
23. Hubbard, A. E. & Mountain, D. C. *Science* **222**, 510-512 (1983).
24. Reuter, G. & Zenner, H. P. *Hearing Res.* **43**, 219-230 (1990).
25. Reuter, G., Gitter, A., Thurm, U. & Zenner, H. P. *Hearing Res.* **60**, 236-246 (1992).
26. Sellick, P. M., Patuzzi, R. B. & Johnstone, B. M. *J. acoust. Soc. Am.* **72**, 131-141 (1982).
27. Ruggero, M. & Rich, N. J. *Neurosci.* **11**, 1057-1067 (1991).
28. Nuttall, A. L., Dolan, D. F. & Avinash, G. *Hearing Res.* **51**, 203-214 (1991).

ACKNOWLEDGEMENTS. We thank P. Kolston, M. Hooley, M. Evans, R. Nobili and J. Gale for critical comments on the manuscript. This work was supported by the Wellcome Trust, the Wolfson Foundation, the EC and the Royal Society.

Inhibition of contractile vacuole function *in vivo* by antibodies against myosin-1

S. K. Doberstein*, I. C. Baines†, G. Wiegand‡, E. D. Korn† & T. D. Pollard*

* Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, Baltimore, Maryland 21225, USA

† Laboratory of Cell Biology, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

‡ Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21225, USA

MYOSIN-I is thought to supply the force for movement of cell membranes relative to actin filaments (reviewed in refs 1, 2), but confirmation of this hypothesis has been difficult because of the presence of multiple isoforms of myosin-I and other unconventional myosins in most cells³. We report here the first evidence that a myosin-I isoform is essential for a specific class of intracellular membrane movements *in vivo*. In *Acanthamoeba*, the contractile vacuole is an autonomous structure which fuses with the plasma membrane to control the water content of the cell. Because myosin-IC is the only myosin-I isoform concentrated in the contractile vacuole complex^{4,5}, and a protein antigenically related to myosin-IC is located on or near the *Dictyostelium* (slime mould) contractile vacuole⁶, we thought this organelle might provide the best opportunity to demonstrate a relationship between myosin-I and membrane motility. Antibodies that inhibit the activity of *Acanthamoeba* myosin-IC *in vitro* interfere with expulsion of excess water by the contractile vacuole *in vivo*, leading to overfilling of this organelle and cell lysis. Myosin-IC may generate the force required to contract the vacuole and may also be involved in transfer of water to the contractile vacuole during refilling.

Antibodies raised against a synthetic peptide with the sequence of the myosin-IC phosphorylation site⁴ inhibited myosin-IC activity *in vitro* in three assays (Table 1): (1) phosphorylation of myosin-IC by the myosin-I heavy chain kinase (phosphorylation is required for activation of *Acanthamoeba* myosin-I ATPase activity^{7,8} and motility *in vitro*^{9,10}); (2) actin-activation of the Mg²⁺-ATPase activity of prephosphorylated myosin-IC; and (3) crosslinking of actin filaments by myosin-IC, indicating that the ATP-sensitive actin binding site (which

TABLE 1 Antibody inhibition of myosin I activities *in vitro*

Experiment	Antibody concentration (μM)	Activity (% of control)		
		ATPase	Crosslinking	Phosphorylation
Myosin-IC + anti-myosin-IC	0.01	100	95	100
	0.06	68	—	100
	0.11	—	0	—
	0.63	0	—	87
	2.50	—	—	41
	3.13	—	—	3
Myosin-IB + anti-myosin-IB	3.65	—	—	0
	0.01	100	84	—
	0.06	32	—	—
	0.11	—	0	—
	0.63	0	0	—

Anti-myosin-IC IgG and anti-myosin-IB IgG were tested for inhibition of three activities: actin-activated Mg²⁺-ATPase activity; crosslinking of F-actin; phosphorylation of myosin-I heavy chain. Myosin-I and myosin-I heavy chain kinase were purified by the method of Lynch *et al.*²⁷. Actin was purified by the method of Spudich and Watt²³. The Mg²⁺-ATPase activity of 0.2 μM myosin-I was assayed as described^{24,25} in the presence and absence of 4 μM F-actin, with the basal activity in the absence of F-actin being subtracted to give the actin-activated Mg²⁺-ATPase activity (control activity: 14.8 s⁻¹). Crosslinking of F-actin (4 μM) by myosin-I (0.2 μM) was measured by a pelleting assay using a low centrifugal force (18,500g) which was just sufficient for pelleting only crosslinked F-actin²⁶. The mixture before pelleting and the supernatant after pelleting were analysed by SDS-PAGE and scanning densitometry to calculate the proportion of myosin-I present as F-actin crosslinks (control value: 100%). Myosin-IC (2 μM) was phosphorylated by autophosphorylated myosin-I heavy chain kinase (60 nM); autophosphorylation is required for kinase activity²⁶. Kinase was autophosphorylated by a 30 min incubation at 30 °C in 2.5 mM ATP, 3.5 mM MgCl₂, 2 mM EGTA, and 50 mM imidazole, pH 7.0. Myosin-I was phosphorylated in the same buffer supplemented with [γ-³²P]ATP (control activity: 1 mol P/mol myosin-I). Relative phosphorylation levels were determined by densitometer scans of autoradiograms after SDS-PAGE.

is near the antigenic epitope in the globular head) was blocked.

We used syringe loading¹¹ to introduce these antibodies into living *Acanthamoeba* cells. We added rhodamine-dextran to the loading solution, and isolated the loaded cells by fluorescence-activated cell sorting (Fig. 1). The combination of these two techniques is considerably easier than microinjection, particularly with *Acanthamoeba*¹², and rapidly yields a large population of loaded cells amenable to statistical analysis. Syringe-loaded anti-myosin-IC IgG was associated with the contractile vacuole

and plasma membrane after fixation of unlysed loaded cells (Fig. 2h).

Acanthamoeba cells loaded with anti-myosin-IC IgG lysed in response to osmotic stress, as judged by microscopic observation (Fig. 3a). Lysis was directly related to osmotic stress, because cells loaded with anti-myosin-IC antibodies and incubated in isotonic growth medium lysed at a much lower frequency than those incubated in water (Fig. 3a). The fraction of cells that lysed directly depended on the fraction of cells above a specific fluorescence level; therefore, lysis occurred when a threshold dose of antibody was introduced into the cell. We determined empirically that a fluorescence level of 700 (arbitrary units) was the minimum required for lysis (Fig. 3a). Assuming the same quantum yield for fluorescence from rhodamine-dextran loaded amoebae and quantitative flow cytometry standards, proportional uptake of IgG and rhodamine-dextran, and the myosin-IC content previously estimated⁵, the threshold fluorescence level was equivalent to an intracellular stoichiometry of about 0.5 IgG binding sites per myosin-IC heavy chain. Control populations loaded with anti-myosin-IB IgG, purified preimmune IgG, bovine serum albumin, or inactivated anti-myosin-IC IgG lysed at low levels and, in each of these cases, the fraction of lysed cells was unrelated to the extent of loading (Fig. 3a). The small percentage of cells that lysed in these control samples may have done so because of damage during the loading procedure irrespective of the amount of protein loaded.

Release of loaded IgG into the supernatant is another way to measure the number of lysed cells (Fig. 3b). The time course of IgG release suggested that all loaded populations contained some damaged cells which lysed within 10 minutes and which accounted for almost all of the lysed cells in the negative control samples. Cells loaded with anti-myosin-IC IgG had a second, larger population that lysed more slowly over 3–4 hours in response to osmotic stress.

DAPI (4',6-diamidino-2-phenylindole) staining combined with flow cytometry is a third method for measuring lysis of

cells loaded with anti-myosin-IC IgG after osmotic shock (Fig. 3c). DAPI is a DNA-binding fluorescent dye that does not cross the plasma membrane of intact cells during short incubations and therefore stains only DNA from lysed cells in this assay. Lysed cells, nuclei, and possibly mitochondria have high DAPI fluorescence and low rhodamine fluorescence, intact cells have low DAPI and high rhodamine fluorescence, and membrane fragments and other organelles have low fluorescence from both DAPI and rhodamine. The largest population of control cells loaded with anti-myosin-IB IgG and osmotically stressed consisted of intact cells (49%, Fig. 3d) that had light-scattering characteristics similar to unloaded intact cells (data not shown). The remaining major groups were composed of lysed cells and nuclei (17%), and small organelles and membrane fragments (29%). The ratio of these three groups was dramatically different for anti-myosin-IC-loaded osmotically shocked cells. Less than 3% of the cells remained intact and the two major populations (Fig. 3c) consisted of: (1) lysed cells, nuclei and possibly mitochondria (49%); and (2) membrane fragments and other organelles (44%).

We also observed cells loaded with anti-myosin-IC antibodies by time-lapse video microscopy after osmotic shock. Lysis in these cells appeared to occur by two distinct mechanisms. Many cells lysed because of overextension of the plasma membrane by the distended contractile vacuole (Fig. 2a–g), with the vacuole occasionally remaining intact even after total rupture of the plasma membrane (Fig. 2d). In another group of cells, the vacuole suddenly stopped increasing in size after contraction, followed by swelling of the cytoplasm and lysis.

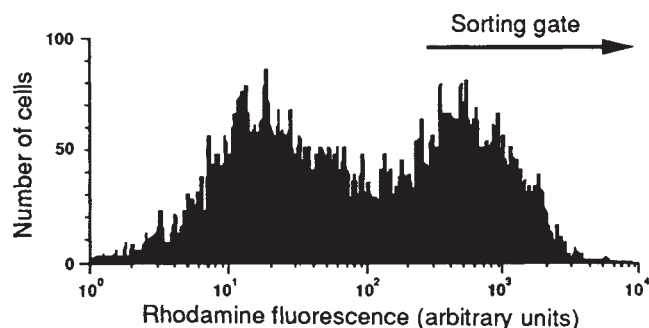


FIG. 1 Fluorescence distribution of cells after loading with immunoglobulins in buffer containing rhodamine-labelled dextran. Note the bimodal distribution of fluorescence. After sorting, no cells with fluorescence below 50 fluorescence units remained in the sample.

METHODS. *Acanthamoeba* cells were grown to stationary phase and concentrated to 10^7 cells ml^{-1} in loading buffer¹² containing 2 mg ml^{-1} purified immunoglobulins and 20 mg ml^{-1} rhodamine-labelled dextran (M_r 73,100; Sigma) before loading. Cells were syringe-loaded¹¹ with five slow passes through a 30-gauge needle. The syringe plunger was drawn back to the 1 ml mark for filling, and the loading suspension was forced out of the syringe at roughly 0.2 ml s^{-1} . The suspension was centrifuged for 1 min at 5,000g, the loading buffer was removed, and the cells were washed briefly in growth medium. The cells were resuspended in 1–4 ml growth medium and sorted on a Becton-Dickinson FACStar fluorescence-activated cell sorter (FACS). The population was screened for viable cells by forward and side light scattering and the live cells were sorted on the basis of rhodamine fluorescence in the same FACS run. Cells with >10 times the median fluorescence of control cells were accepted as loaded.

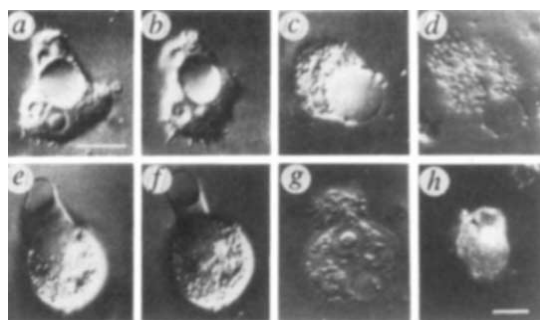
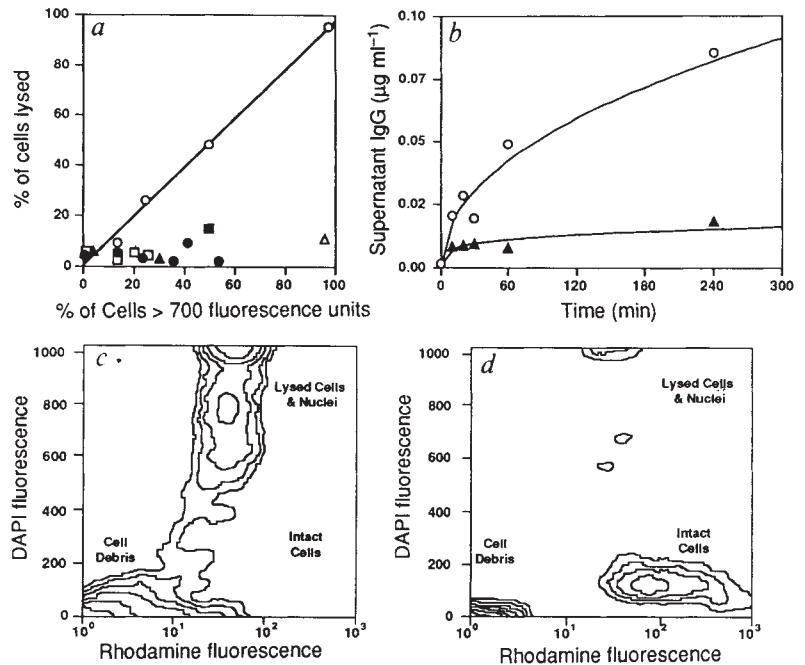


FIG. 2 Video microscopy of cell lysis. a–d, Time course of lysis of a single cell over 35 min (a, 3 min 30 s; b, 17 min 20 s; c, 30 min 5 s; d, 35 min). The plasma membrane ruptured (c), whereas the contractile vacuole remained substantially intact. e–g, Time course of lysis of a single cell over 12 min (e, 1 min 15 s; f, 7 min 15 s; g, 12 min 5 s). The enlarged contractile vacuole distended the plasma membrane and ruptured, followed by lysis of the cell. Scale bar in a, 15 μm . h, Immunofluorescence detection of anti-myosin-IC IgG in loaded cells fixed after incubation in growth medium and stained with fluorescein conjugated goat anti-rabbit IgG. The anti-myosin-IC antibody labels the contractile vacuole complex (arrowhead) and plasma membrane. Scale bar, 10 μm . **METHODS.** Cells were syringe-loaded with anti-myosin-IC IgG and sorted by FACS (see Fig. 1). After sorting, the cells were resuspended in 250 μl growth media and pipetted onto a glass coverslip. After 15 min for the cells to attach, the excess media was wicked off and the cover slip was inverted onto a glass microscope slide with a drop of distilled water within a ring of petroleum jelly. The cells were observed on a Zeiss Axiocvert 135TV microscope equipped with a Plan-Neofluar $\times 100$ objective using differential interference contrast optics (courtesy of D. B. Murphy). The image was collected with a Hamamatsu C2400 television camera and controller and was processed with a Hamamatsu Argus-10 image processor. The processed images were photographed using a Polaroid freeze frame video recorder and 35 mm camera adapter with Kodak TMax 100 film. For immunofluorescence, loaded cells without FACS sorting were allowed to settle in growth medium onto poly-L-lysine coated microscope slides for 20 min before fixation, permeabilization in 100% acetone and reduction with sodium borohydride as described^{4,5}. Anti-myosin-IC IgG was detected by goat anti-rabbit IgG conjugated to fluorescein (10 $\mu\text{g ml}^{-1}$; Molecular Probes).

FIG. 3 *Acanthamoeba* cells loaded with antibodies against myosin-IC lyse when incubated in distilled water. **a**, Percentage of cells that lyse within 2 h of incubation as a function of fluorescence level. $\circ$, Cells loaded with anti-myosin-IC IgG and incubated in water; $\blacksquare$, cells loaded with anti-myosin-IC IgG and incubated in growth medium; $\blacktriangle$, cells loaded with preimmune IgG and incubated in water; $\triangle$, cells loaded with anti-myosin-IB IgG and incubated in water; $\square$, cells loaded with inactivated anti-myosin-IC IgG and incubated in water; $\bullet$, cells loaded with BSA and incubated in water. **b**, Time course of the release of IgG from lysing cells. $\circ$, Cells loaded with anti-myosin-IC IgG and incubated in water; $\blacktriangle$, cells loaded with preimmune IgG and incubated in water. **c, d**, Flow cytometry analysis of cell lysis. As cells lyse, they lose the rhodamine fluorescence signal and increase in DAPI fluorescence. As analysed by light scattering, the population with high rhodamine fluorescence is primarily made up of intact cells, whereas the high DAPI/low rhodamine population is a mixture of lysed cells and free nuclei (and perhaps mitochondria) of lysed cells. The population with both low DAPI and low rhodamine signals is made up of small cell fragments with no DNA content. **c**, Cells loaded with anti-myosin-IC IgG and incubated in water. **d**, Cells loaded with anti-myosin-IB IgG and incubated in water.

METHODS. Cells were syringe-loaded and sorted by FACS (see Fig. 1). The distribution of fluorescence level in the cell population was determined by flow cytometry during sorting, and the percentage of cells above 700 arbitrary fluorescence units was calculated. Immediately after sorting, the cells were allowed to recover in 1 ml growth medium for at least 30 min. The data shown in **a** were obtained by incubating the cells for 15 min on tissue culture slides (Miles Scientific, Naperville, IL) in growth medium to allow the cells to attach to the surface. The growth medium was gently removed and either distilled water or additional growth medium was added. After 2 h at least 500 cells in each sample were examined by each of two observers, and phase transparent cells were scored as lysed. The flow cytometer was calibrated with fluorescent standards of known rhodamine content. For data shown in **b**, loaded and sorted cells were allowed to attach to wells in a 16-well tissue culture plate for 15 min. The growth medium was gently removed and the cells covered with



water. The supernatant was sampled at various time points, and IgG was quantified by ELISA. For data in **c, d**, loaded and sorted cells were allowed to recover in growth medium, and then collected by centrifugation for 5 min at 1,000g and resuspended in water. After a 40 min incubation, DAPI (4',6-diamidino-2-phenylindole, Molecular Probes) was added to the suspension, which was incubated for 5 min, and then analysed by dual wavelength flow cytometry. DAPI undergoes a 20-fold increase in fluorescence on binding to DNA, and does not cross intact cell membranes during short incubations. Anti-myosin-IC and -IB polyclonal antibodies were previously described^{4,5}. Inactivated antibodies were biotinylated with four to five moles of biotin per mole of antibody. The biotinylated IgG did not recognize myosin-I on protein immunoblots, nor did it stain cells by indirect immuno-fluorescence.

Contraction of the vacuole was severely inhibited in cells loaded with anti-myosin-IC IgG, but not those loaded with anti-myosin-IB IgG. We observed at least 100 vacuole filling and contraction cycles in both loaded and unloaded cells. The average time between contractions after osmotic shock was increased from 92 s for unloaded cells and 105 s for cells loaded with anti-myosin-IB, to 237 s for cells loaded with anti-myosin-IC. Abortive partial contractions accounted for 18% of the observed contractions in anti-myosin-IC-loaded cells; no partial contractions were observed in either unloaded cells or anti-myosin-IB-loaded cells. Contraction of isolated contractile vacuoles *in vitro* requires ATP (data not shown).

These observations strongly suggest that myosin-IC is essential for active contraction of the vacuole. The exact mechanism remains unclear, but it is apparent that it is an active process involving an actomyosin-I interaction, rather than a passive mechanism of membrane rupture and concomitant contractile vacuole collapse¹³. Myosin-IC may also be a motor for movement of spongiome vesicles during filling of the vacuole. It remains a possibility that anti-myosin-IC IgG disrupts water regulation at the plasma membrane, leading to overloading of the contractile vacuole system, but in that case we would expect that the time between vacuole contractions would decrease before cell lysis rather than dramatically increase as we observed. Also, this possibility does not explain the presence of 'frozen' non-contracting contractile vacuoles observed only in anti-myosin-IC-loaded cells.

Our observation that osmotic shock is required to reveal the phenotype of myosin-IC deficient cells has more general implications. The actomyosin system in motile cells is quite robust under

laboratory culture conditions, with cells surviving even after inhibition or genetic elimination of major cytoskeletal proteins^{12,14-22}. Subjecting these cells to severe stress analogous to osmotic shock might reveal interesting new phenotypes. $\square$

Received 10 May; accepted 29 July 1993.

- Pollard, T. D., Doberstein, S. K. & Zot, H. G. *Rev. Physiol.* **53**, 653-681 (1991).
- Hammer, J. A. III *Trends Cell Biol.* **1**, 50-56 (1991).
- Cheney, R. E. & Mooseker, M. S. *Curr. Opin. Cell Biol.* **4**, 27-35 (1992).
- Baines, I. C. & Korn, E. D. *J. Cell Biol.* **111**, 1895-1904 (1990).
- Baines, I. C., Brzeska, H. & Korn, E. D. *J. Cell Biol.* **119**, 1193-1203 (1992).
- Zhu, Q. & Clarke, M. J. *Cell Biol.* **118**, 347-358 (1992).
- Maruta, H. & Korn, E. D. *J. Biol. Chem.* **252**, 8329-8332 (1977).
- Lynch, T. J., Brzeska, H., Miyata, H. & Korn, E. D. *J. Biol. Chem.* **264**, 19333-19339 (1989).
- Albanesi, J. P. et al. *J. Biol. Chem.* **260**, 8649-8652 (1985).
- Zot, H. G., Doberstein, S. K. & Pollard, T. D. *J. Cell Biol.* **116**, 367-376 (1991).
- Clarke, M. S. F. & McNeil, P. L. *J. Cell Sci.* **102**, 533-541 (1992).
- Sinard, J. H. & Pollard, T. D. *Cell Motil. Cytoskel.* **12**, 42-53 (1989).
- Wigg, D., Bovee, E. C. & Jahn, T. L. *J. Protozool.* **14**, 104-108 (1967).
- Titus, M. A., Wessels, D., Spudich, J. A. & Soll, D. *Molec. Biol. Cell* **4**, 233-246 (1993).
- Wessels, D., Murray, J., Jung, G., Hammer, J. A. & Soll, D. R. *Cell Motil. Cytoskel.* **20**, 301-315 (1991).
- Jung, G. & Hammer, J. H. III *J. Cell Biol.* **110**, 1955-1964 (1990).
- Witte, W., Schleicher, M. & Noegel, A. A. *Cell* **68**, 53-62 (1992).
- Cox, D. J. *Cell Biol.* **116**, 943-955 (1992).
- DeLozanne, A. & Spudich, J. A. *Science* **236**, 1086-1091 (1987).
- Knecht, D. A. & Loomis, W. F. *Science* **236**, 1081-1086 (1987).
- Wessels, D. et al. *Dev. Biol.* **128**, 164-177 (1988).
- Furukawa, R., Butz, S., Fleischmann, E. & Fechtmeier, M. *Protoplasma* **169**, 18-27 (1992).
- Spudich, J. A. & Watt, S. J. *J. Biol. Chem.* **246**, 4866-4871 (1971).
- Pollard, T. D. & Korn, E. D. *J. Biol. Chem.* **248**, 4682-4690 (1973).
- Albanesi, J. P., Lynch, T. J., Fujisaki, H. & Korn, E. D. *J. Biol. Chem.* **261**, 10445-10449 (1986).
- Brzeska, H., Lynch, T. J. & Korn, E. D. *J. Biol. Chem.* **265**, 3591-3594 (1990).
- Lynch, T. J., Brzeska, H., Baines, I. C. & Korn, E. D. *Meth. Enzym.* **196**, 12-23 (1990).

ACKNOWLEDGEMENTS. We thank D. B. Murphy and C. A. Ritter for assistance with video microscopy; and S. Atkinson, E. De La Cruz, K. Ryan and J. A. Hammer III for reading the manuscript. This work was funded in part by a NIH grant to T.D.P.

Cloning and characterization of the vasopressin-regulated urea transporter

Guofeng You, Craig P. Smith, Yoshikatsu Kanai, Wen-Sen Lee, Matthias Stelzner & Matthias A. Hediger*

Renal Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA

UREA is the principal end product of nitrogen metabolism in mammals¹. Movement of urea across cell membranes was originally thought to occur by lipid-phase permeation, but recent studies have revealed the existence of specialized transporters with a low affinity for urea ($K_m > 200$ mM)². Here we report the isolation of a complementary DNA from rabbit renal medulla that encodes a 397-amino-acid membrane glycoprotein, UT2, with the functional characteristics of the vasopressin-sensitive urea transporter previously described in *in vitro*-perfused inner medullary collecting ducts^{3,4}. UT2 is not homologous to any known protein and displays a unique pattern of hydrophobicity. Because of the central role of this transporter in fluid balance^{1,3-7} and nitrogen metabolism⁸, the study of this protein will provide important insights into the urinary concentrating mechanism and nitrogen balance.

Using expression-cloning in *Xenopus* oocytes^{9,10}, we isolated a 3,100-base-pair (bp) cDNA from a rabbit kidney medullary cDNA library, which induced uptake of ¹⁴C-urea 10- to 17-fold above that of water-injected controls (Fig. 1a). The clone, named UT2, contains an open reading frame that predicts a 397-residue protein (Fig. 2a). A search of translated sequence databases revealed that UT2 is not homologous to any known sequence.

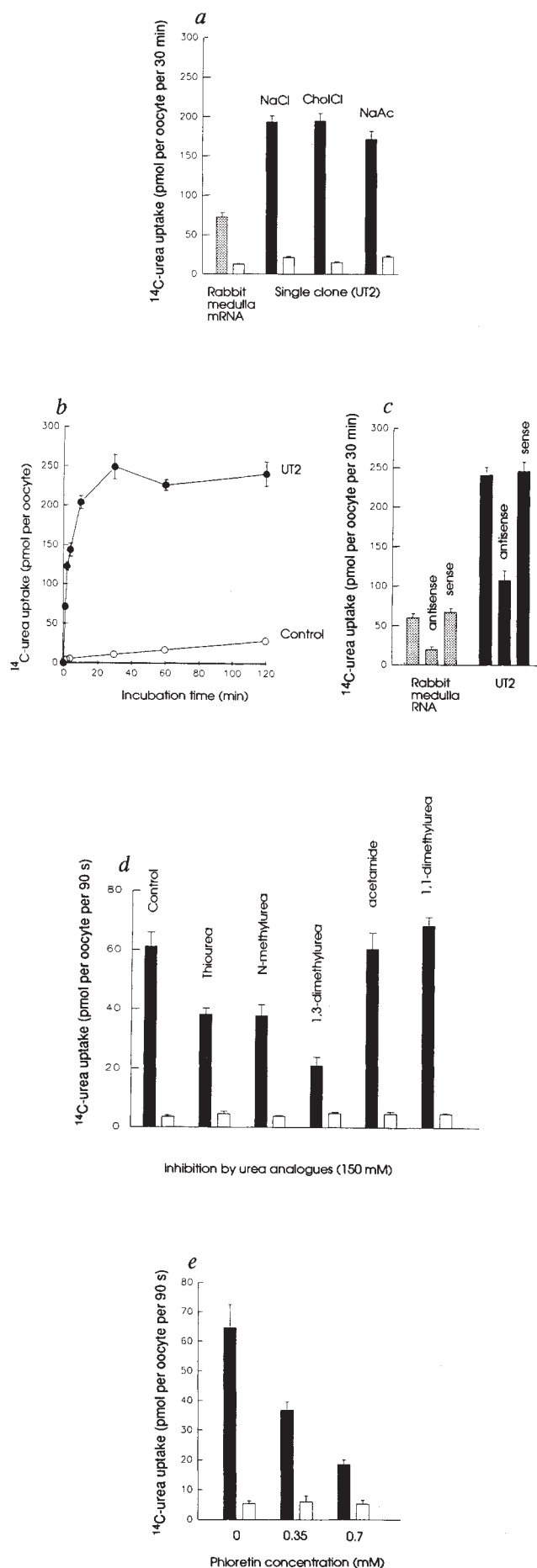
Urea transport in the mammalian kidney is generally considered to be passive¹¹, although there is some evidence for active transport¹². Replacing NaCl with choline chloride or sodium acetate did not alter UT2-induced urea transport in oocytes, indicating that transport is not Na⁺- or Cl⁻-dependent (Fig. 1a). Furthermore, bath-applied urea (1–200 mM) did not evoke a significant current in UT2 complementary RNA-injected oocytes subjected to two-microelectrode voltage-clamp analysis¹⁰ (data not shown). This is consistent with transport of urea by UT2 being electroneutral.

Uptake of ¹⁴C-urea (1 mM) into UT2-injected oocytes reached a maximum of ~250 pmol per oocyte in 10 to 30 min (Fig. 1b), which corresponds to an intracellular urea concentration of 1 mM if the effective cytoplasmic volume of an oocyte is assumed to be 0.25 μ l. Thus ¹⁴C-urea had almost completely equilibrated across the oocyte plasma membrane after 10 min suggesting that transport of urea by UT2 is by facilitated transport.

The calculated urea permeability of UT2-injected oocytes based on initial rates of uptake is 4.5×10^{-5} cm s⁻¹, in agreement with the value of 11.6 – 13.1×10^{-5} cm s⁻¹ for *in vitro*-perfused terminal inner medullary collecting ducts (IMDC)¹³. As the urea permeability of water-injected control oocytes is 1.9×10^{-6} cm s⁻¹, UT2 increased the permeability of oocytes to urea 23-fold.

Depletion of kidney medulla messenger RNA before injection into oocytes with an antisense oligonucleotide corresponding to sequences near the UT2 start codon resulted in a 67% reduction of ¹⁴C-urea uptake (Fig. 1c). This indicates that UT2 represents the major urea transporter in the IMCD.

Inhibition of urea uptake (1 mM) by urea analogues (at 150 mM) showed that, of the analogues tested, 1,3-dimethylurea



* To whom correspondence should be addressed at Brigham and Women's Hospital.

effectively inhibited urea transport by 65%, thiourea and *N*-methylurea inhibited it by 38%, and acetamide and 1,1-dimethylurea had no significant effect (Fig. 1d). Furthermore phloretin at 0.35 mM or 0.7 mM inhibited UT2-mediated uptake of 1 mM urea by 48 and 78%, respectively (Fig. 1e). This pattern of inhibition is almost identical to that of apical vasopressin-sensitive urea transport reported for *in vitro*-perfused rat kidney IMCD^{14,15}.

Saturation of transport with increasing substrate concentration is a characteristic feature of carrier-mediated transport. The maximal urea concentration that can be used with oocytes is ~200 mM but urea transport in the perfused IMCD and erythrocytes has K_m values for urea significantly above 200 mM^{2,16,17}. Our results show that there is no saturation of UT2-mediated urea transport in the concentration range between 1 and 200 mM urea (data not shown).

Hydropathy analysis of UT2 (Fig. 2b) revealed two large extended hydrophobic domains (A and C) (Fig. 2b), unlike previously studied transporters which have several distinct hydrophobic transmembrane segments interspersed with hydrophilic regions^{9,18}.

In vitro translation of UT2 cRNA using rabbit reticulocyte lysates followed by SDS-polyacrylamide gel electrophoresis yielded a protein with an apparent M_r of 40K in the absence of, and of 45K in the presence of dog pancreatic microsomes. This 5K shift in M_r was reversed by treatment with endoglycosidase-H and is indicative of *N*-glycosylation at a single site¹⁹. This site is probably Asn 210 as the only other potential *N*-glycosylation site, Asn 288, is part of a hydrophobic domain. Domain B, which includes Asn 210, is therefore likely to be extracellular. A

putative model for UT2 is shown in Fig. 2c. It is conceivable, however, that owing to the presence of only a few charged residues in domains A and C (Fig. 2b) a major portion of the protein is entirely embedded in the membrane.

UT2 has three potential phosphorylation sites, two sites for protein kinase A (at Ser 31 and Ser 386) and one for protein kinase C (at Ser 13) (Fig. 2c). Vasopressin has been proposed

a

```

1  MEDSSEIKVETASSRTSWIQSSVAAGGKRISRALGYITGEMKECAEGLKD
20
60  KSPVFQFLDWVLRGTQVMFVNPLSGILIVIGLVQNPWWAAGCLGTV
80
120  MSTLTALILSQDRSAIASGLHGNGVLVGLLIATVFSDRGDYWWLLLPVI
140
160  VMGMSCPILSSALGTIFSKWDLVFETLPENIAVTLYLAATHYNNFFPTT
180
220  LLQPVSSVPTNINSETQVPLLRRAIPVGIGQVYGCNPWTGGIFLIALFI
240
260  SSPLICLHAAIGSTMGLAALTATIPDSIYFGLCGFNSTLACIIVGGMF
300
320  YVITWQTHLLAVACALFAAYVGAALTNVLSVFLPTCTWPPFCISALIFLL
340
360  LTTNNPAIYKLPISKVTYPEANRTYYLTQERNRRSSITIKYQAYDVVS*
380
397

```

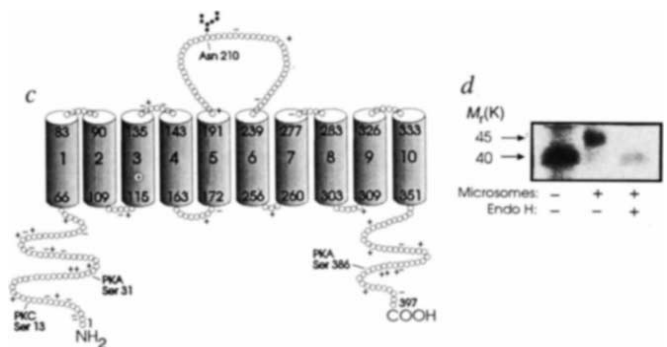
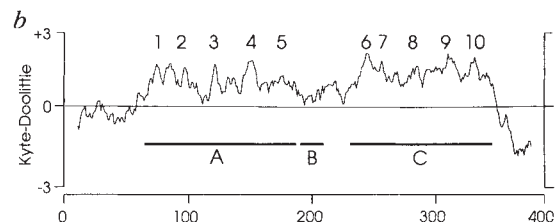


FIG. 2 a, Amino-acid sequence of UT2. An open reading frame of UT2 cDNA (3.1 kb) was identified from nucleotides 1,025–2,215 that predicts a 397-residue membrane protein. The start codon is preceded by multiple stop codons and corresponds to a Kozak initiation of translation site (CCCATGG)²⁸. Hydrophobic regions predicted to be membrane-spanning (18–21 residues long) are indicated by lines numbered 1–10. The potential *N*-glycosylation sites of the type N-X-T/S are boxed (Asn 210 and Asn 288). b, Kyte–Doolittle hydropathy analysis using a window of 21 (ref. 29). Potential membrane-spanning regions are numbered 1–10. A and C denote extended hydrophobic regions, and B is a more hydrophilic region predicted to be extracellular. c, A membrane model of UT2 protein. Putative membrane-spanning regions were predicted using the algorithm of Eisenberg³⁰. Arg, Lys and His residues are indicated by plus signs and Glu and Asp by minus signs. Potential protein kinase A (PKA) and protein kinase C (PKC) phosphorylation sites are indicated if predicted to be cytoplasmatic. d, *In vitro* translation of UT2 cRNA. Translation is shown in the absence of pancreatic microsomes (lane 1), in the presence of microsomes after centrifugation (lane 2) and after deglycosylation with endoglycosidase-H (lane 3), as analysed by 15% SDS-PAGE.

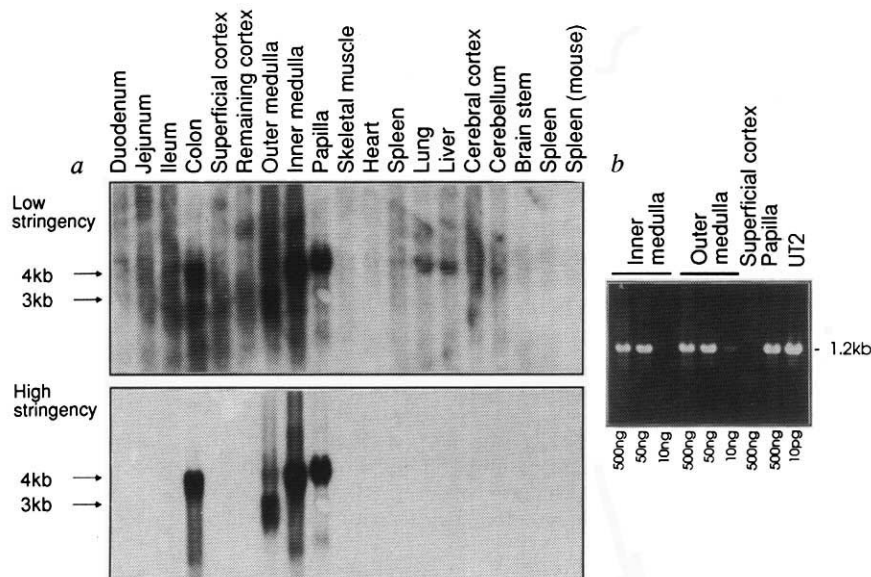
METHODS. Rabbit kidney medulla poly(A)⁺RNA was size-fractionated by preparative agarose gel electrophoresis³¹. Fractions were analysed by expression in oocytes. A 3–4.5-kb fraction was enriched for mRNA that induced urea transport (not shown). This fraction was used to construct a directional cDNA library. The procedures for expression cloning, sequencing and *in vitro* translation have been described^{9,10,19}.

FIG. 1 Expression of UT2 in *Xenopus* oocytes. Filled bars and circles represent oocytes injected with UT2 cRNA; open bars and circles, water-injected oocytes; and cross-hatched bars, rabbit kidney medulla poly(A)⁺RNA-injected oocytes. Columns represent the mean $\pm$ s.e.m. ($n=5-8$ oocytes). a, Ion dependency of ¹⁴C-urea uptake into *Xenopus* oocytes. b, Time course of the UT2-mediated urea uptake into oocytes. c, Hybrid depletion of rabbit kidney medulla poly(A)⁺RNA by incubation with an antisense UT2 oligonucleotide. Residual uptake was probably due to incomplete depletion, as there was only a 55% reduction in uptake after depletion of UT2 cRNA with the UT2 antisense oligonucleotide. Bars represent uptakes after subtraction of water-injected controls. d, Inhibition of UT2-mediated urea uptake by urea analogues. Uptake of ¹⁴C-urea (1 mM) in the presence of 150 mM of the urea analogues thiourea, *N*-methylurea, 1,3-dimethylurea, acetamide or 1,1-dimethylurea and 50 mM mannitol. Inhibition of urea uptake by urea analogues was studied under pre-equilibrium conditions (90 s uptakes). e, Phloretin inhibition.

METHODS. *Xenopus laevis* oocyte expression was studied as described^{9,10}. After cRNA injection, oocytes were incubated at 18 °C for 3 d, after which uptake of ¹⁴C-urea was measured. Oocytes were preincubated in 200 mM mannitol, 2 mM KCl, 1 mM MgCl₂, 1 mM CaCl₂, 10 mM HEPES buffer, 5 mM Tris, pH 7.4, for 1 hour. For the measurement of urea uptake, 2.74 μ Ci ¹⁴C-urea ml⁻¹ and 1 mM urea were added to the preincubation solution. Unlabelled urea was first deionized by passing through an ion-exchange column (AG501-X8(D), 20–50 mesh; BioRad). After uptake, oocytes were washed with ice-cold uptake solution containing 1 mM unlabelled urea, dissolved in 10% SDS, and the radioactivity counted in a scintillation counter. In ion-dependency experiments (a), mannitol was replaced by NaCl (100 mM), choline chloride (100 mM) or sodium acetate (100 mM). Urea uptake was measured during 30 min. For hybrid depletion studies, rabbit kidney medulla poly(A)⁺RNA or UT2 cRNA (0.9 μ g μ l⁻¹, heat-denatured) was incubated with antisense or sense oligonucleotides (0.3 μ g μ l⁻¹), corresponding to a region near the start codon (nucleotides 1,246–1,269) in the presence of 50 mM NaCl at 42 °C for 20 min, cooled on ice and injected into oocytes²⁷. The uptake of ¹⁴C-urea was measured during 30 min. For inhibition studies, all analogues were deionized as already described and ¹⁴C-urea flux was measured during 90 s. For measuring phloretin inhibition, 0.5 M phloretin in ethanol, or ethanol at the same concentration for controls, was added to the uptake solutions. Oocytes were preincubated for 15 min in the phloretin-containing solution and ¹⁴C-urea flux was measured during 90 s.

FIG. 3 a, Expression of UT2 in different tissues. Low- and high-stringency northern blots of poly(A)⁺ RNA (3 µg) from rabbit tissues probed with ³²P-labelled UT2 cDNA. A strong 4-kb band was detected in colon, kidney inner medulla, papilla and a 3-kb band in the outer medulla. At low stringency, the probe also hybridized weakly to 4-kb bands in RNA samples from liver and lung. Total RNA (10 µg) was used for papilla. The last two lanes show spleen RNA from anaemic (acetylphenyl hydrazine-treated) rabbits or mice. No signal was detected in these lanes, indicating that the erythrocyte urea transporter is different from UT2. The difference between the message sizes in kidney inner and outer medulla may indicate different polyadenylation sites. **b**, Detection of UT2 expression by PCR. Amplification of the entire UT2 coding region from reverse-transcribed inner medulla, outer medulla and papilla mRNA and from UT2 cDNA shows a 1.2-kb amplification product for all samples, except for the negative control tissue (superficial cortex). The dose dependency (10–500 ng RNA) of the PCR reactions indicated that the product from the outer medulla is not the result of PCR amplification of the small amount of 4-kb message. Digestion of the PCR products with the enzymes *Bam*HI, *Pst*II, *Bgl*II and *Sst*I, which targeted internal cleavage sites of the UT2 coding region, gave identical cleavage patterns for outer medulla, inner medulla and UT2 (not shown), indicating that the coding regions of the 3- and 4-kb messages and of UT2 cDNA are identical.

METHODS. Northern blot analysis and PCR amplification were done



as described¹⁰. The northern blot filter was hybridized at 35 °C (low stringency) and 42 °C (high stringency) in 50% formamide using full-length UT2 as a probe, labelled with ³²P using the T7 QuickPrimer Kit (Pharmacia). The filter was washed in 0.1 × SSC, 0.1% SDS at 42 °C (low stringency) and 65 °C (high stringency). RNA from spleen of rabbit and mouse and treated with acetylphenyl hydrazine was provided by S. Alper.

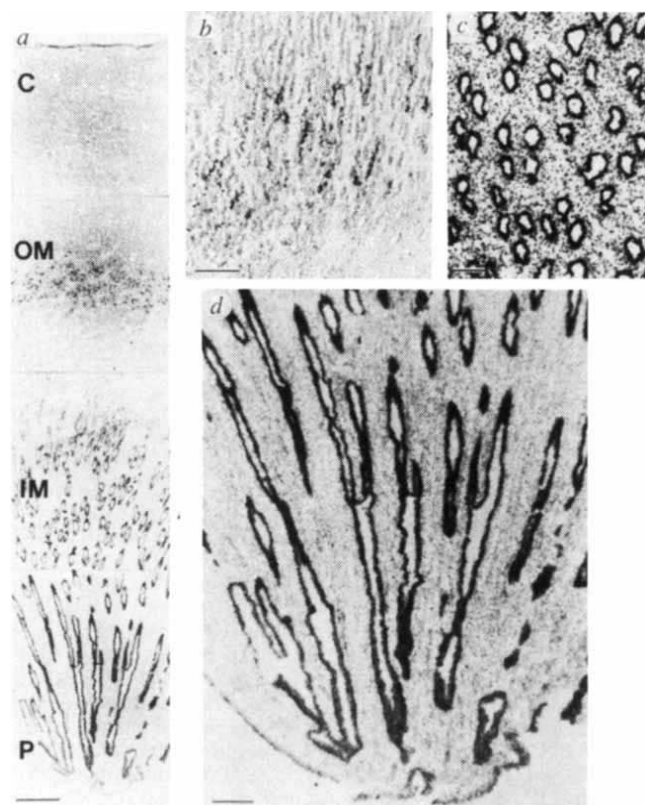


FIG. 4 Localization of UT2 mRNA by *in situ* hybridization of rabbit kidney. **a**, Parasagittal section of rabbit kidney showing a strong signal for UT2 cRNA probe in the terminal part of the IMCD and the papilla. A specific signal was also seen in the inner stripe of the outer medulla. Scale bar, 500 µm. Abbreviations: C, cortex; OM, outer medulla; IM, inner medulla; P, papilla. **b**, Outer medulla viewed from a plane tangential to the corresponding region in **a**, showing a close-up of the signal in this region. Scale bar, 200 µm. **c**, Cross-section of IMCD. Scale bar, 100 µm. **d**, Close-up of the papillary region shown in **a**. Junctions between converging collecting ducts in the inner medulla are clearly visible. Scale bar, 200 µm.

METHODS. *In situ* hybridization of rabbit kidney was done as described³² using 4% paraformaldehyde-fixed tissue sections (7 µm). Briefly, ³⁵S-labelled sense and antisense RNA probes were synthesized from the full-length clone (in pSPORT) after linearization of plasmid DNA with *Hind*III or *Kpn*I, using T7 or SP6 RNA polymerase, respectively. RNA probes were hydrolysed for 50 min to form probes of ~100 nucleotides.

to stimulate urea transport by means of cyclic AMP-dependent phosphorylation^{14,20}. Oocytes are a suitable system for studying the regulation of membrane transport by direct cAMP-dependent phosphorylation²¹. The cAMP analogues Sp-cAMP-S, dibutyryl-cAMP or 8-bromo-cAMP (30 min preincubation, 1 mM analogues), however, were found to have no effect on

UT2-mediated urea transport (data not shown). Therefore stimulation may not involve direct phosphorylation of UT2 but rather the recruitment of transporters from vesicle pools, as in the case for the renal water channel²².

In high-stringency northern blot analysis, prominent RNA bands of 4 kilobases (kb) were seen in samples from rabbit renal papillary tip, renal inner medulla and colon (Fig. 3a). A band at 3 kb in the outer medulla was also present. Amplification by polymerase chain reaction (PCR) of RNA from inner and outer medulla, with primers from the complete coding region of UT2, gave products of the same size (Fig. 3b), with identical restriction maps (data not shown). This suggests that the UT2 message in the outer medulla has the same coding region as in the inner medulla.

UT2 mRNA in colon appears to be identical to that detected in kidney inner medulla. Urea formed in the liver is probably disposed of not only by excretion in the urine but also by secretion into the colon, where it is hydrolysed by gut microflora⁸. At times of protein restriction, ammonia formed in this way may be salvaged and transported to the liver in the portal vein for further metabolism⁸.

Under low-stringency conditions, the UT2 probe hybridized to distinct 4-kb bands in liver and lung RNA samples. A recent study²³ suggested the existence of a phloretin-inhibitable urea transporter in the mammalian liver but *in vivo* studies have failed to reveal saturable urea transport in lung²⁴. Our finding, however, suggests that a UT2-related urea transporter is present in lung.

In erythrocytes, low-affinity urea transporters control water efflux during the passage through regions of 0.5–0.6 M urea in the kidney¹⁶. We were unable to detect a signal in RNA from spleens of anaemic rabbits or mice (Fig. 3a), suggesting that the urea transporter of the red blood cell is significantly different from UT2.

In situ hybridization showed a strong UT2 signal in epithelial

cells lining the IMCD (Fig. 4). This localization is consistent with the predicted distribution of the vasopressin-regulated urea transporter^{4,5,13,25}. Hybridization in the inner stripe of the outer medulla is congruent with the signal at 3 kb in the outer medulla, as detected by northern analysis (Fig. 3a). This signal may represent a urea transporter located either in the thin descending limbs of the short loops of Henle²⁶ or in descending vasa recta³³ because urea permeability has been shown to be relatively high in these structures. If this proves to be true then UT2 is likely to be involved in urea recycling in this part of the kidney.

In summary, we have cloned and characterized a novel and unique protein which is responsible for carrier-mediated urea transport. □

Received 24 May; accepted 31 August 1993.

1. Marsh, D. J. & Knepper, M. A. in *Handbook of Physiology* 8, *Renal Physiology* Vol. 2 (ed. Windhager, E. E.) 1317–1347 (Oxford Univ. Press, Oxford, 1992).
2. Chou, C.-L., Sands, J. M., Nonoguchi, H. & Knepper, M. A. *Am. J. Physiol.* **258**, F486–F494 (1990).
3. Knepper, M. A. & Star, R. A. *Am. J. Physiol.* **259**, F393–F401 (1990).
4. Sands, J. M., Nonoguchi, H. & Knepper, M. A. *Am. J. Physiol.* **253**, F823–F832 (1987).
5. Knepper, M. A., Sands, J. M. & Chou, C.-L. *Am. J. Physiol.* **256**, F610–F621 (1989).
6. Zhang, R. & Verkman, A. S. *J. Membr. Biol.* **117**, 253–261 (1990).
7. Sands, J. M. & Kokko, J. P. *Kidney Int.* **38**, 695–699 (1990).
8. Langran, M., Moran, B. J., Murphy, J. L. & Jackson, A. A. *Clin. Science* **82**, 191–198 (1992).
9. Hediger, M. A., Coady, M. J., Ikeda, T. S. & Wright, E. M. *Nature* **330**, 379–381 (1987).
10. Kanai, Y. & Hediger, M. A. *Nature* **360**, 467–471 (1992).
11. Knepper, M. A. & Roch-Ramel, F. *Kidney Int.* **31**, 629–633 (1987).
12. Kawamura, S. & Kokko, J. P. *J. clin. Invest.* **58**, 604–612 (1976).
13. Sands, J. M. & Knepper, M. A. *J. clin. Invest.* **79**, 138–147 (1987).
14. Chou, C.-L. & Knepper, M. A. *Am. J. Physiol.* **257**, F359–F365 (1989).
15. Mayrand, R. R. & Levitt, D. G. *J. gen. Physiol.* **81**, 221–237 (1983).
16. Brahm, J. J. *gen. Physiol.* **82**, 1–23 (1983).
17. Toon, M. R. & Solomon, A. K. *Biochim. biophys. Acta* **940**, 266–274 (1988).
18. Mueckler, M. et al. *Science* **229**, 941–945 (1985).
19. Hediger, M. A., Mendlein, J., Lee, H. S. & Wright, E. M. *Biochim. biophys. Acta* **1064**, 360–364 (1991).

20. Star, R. A., Nonoguchi, H., Balaban, R. & Knepper, M. A. *J. clin. Invest.* **81**, 1879–1888 (1988).
21. Uezono, Y. et al. *Recept. Channels* **1**, 233–241 (1991).
22. Fushimi, K. et al. *Nature* **361**, 549–552 (1993).
23. Effros, R. M., Murphy, C., Hacker, A. & Ozker, K. *FASEB J.* **6**, A2073 (1992).
24. Effros, R. M., Murphy, C., Ozker, K. & Hacker, A. *Am. J. Physiol.* **263**, L619–L626 (1992).
25. Sands, J. M., Nonoguchi, H. & Knepper, M. A. in *Vasopressin: Cellular and Integrative Functions* (eds Cowley, A. W. Jr, Liard, J.-F. & Ausiello, D. A.) 137–142 (Raven, New York, 1988).
26. Imai, M., Hayashi, M. & Araki, M. *Pflügers Arch.* **402**, 385–392 (1984).
27. Snutch, T. P., Leonard, J. P., Gilbert, M. M., Lester, H. A. & Davidson, N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3391–3395 (1990).
28. Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
29. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
30. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. J. *J. molec. Biol.* **179**, 125–142 (1984).
31. Hediger, M. A., Tyson, I., Coady, M., Gundersen, C. B. & Wright, E. M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2634–2637 (1987).
32. Kanai, Y. et al. *Am. J. Physiol.* **263**, F1087–F1093 (1992).
33. Pallone, T. L., Work, J., Myers, R. L. & Jamison, R. L. *J. clin. Invest.* (in the press.)

ACKNOWLEDGEMENTS. We thank M. Knepper for discussion, P. Boutin for technical assistance, and A. Yu for help with PCR amplification. This work was supported by an NIH grant to M.A.H. and by an American Heart Association Massachusetts Affiliate Fellowship to G.-F.Y.

Identification of a proline residue as a transduction element involved in voltage gating of gap junctions

Thomas M. Suchyna, Lie Xian Xu*, Feng Gao, Charles R. Fournier & Bruce J. Nicholson†

State University of New York at Buffalo, Department of Biological Sciences, Buffalo, New York 14260, USA

GAP junction channels are structurally distinct from other ion channels in that they comprise two hemichannels which interact head-to-head to form an aqueous channel between cells. Intercellular voltage differences together with increased intracellular concentrations of H^+ and Ca^{2+} cause closure of these normally patent channels¹. The relative sensitivity to voltage varies with the subunit (connexin) composition of the channels². The third of four transmembrane-spanning regions (M3) in connexins has been proposed to form the channel lining³, and a global 'tilting' of the hemichannel subunits has been correlated with channel closure⁴. But specific components involved in transduction of channel gating events have not been identified in either gap junctions or other ion channel classes (however, see model in ref. 5). We have examined a strictly conserved proline centrally located in M2 of connexin proteins. Mutation of this proline (Pro 87) in connexin 26 causes a reversal in the voltage-gating response when the mutant hemichannel is paired with wild-type connexin 26 in the *Xenopus* oocyte system. This suggests that the unique properties associated with this residue are critical to the transduction of voltage gating in these channels.

In soluble proteins, proline is frequently associated with turns in the backbone. Rotation about its N–Ca bond is restricted⁶, creating a fixed ϕ angle which disrupts or distorts regular α -helix or β -sheet structures, inducing a 15–20° bend in the former^{7–9}. This disruption also causes the loss of one or two backbone hydrogen bonds, including that involving the proline amino group⁶. The energy barrier to *cis/trans* isomerizations about the peptide bond preceding proline is also reduced (13 kcal mol⁻¹) compared with other residues (20 kcal mol⁻¹)¹⁰, raising the possibility that proline could form a 'hinge-like' region. Despite these properties, proline has been seen with surprising frequency in presumed helical or β -sheet transmembrane regions of membrane transporters and channels¹¹, including all members of the connexin (Fig. 1a), and ligand-gated channel families (Fig. 1b). Mutagenic studies on transporters have not consistently demonstrated a role for these prolines in the transport process^{12–14}, but how they effect channel function has yet to be tested.

Connexin 26 (Cx26) channels, like most gap junction channels expressed in *Xenopus* oocytes, close symmetrically in response to transjunctional voltage (V_j) of either polarity¹⁵. A minor response to inside-outside voltage (V_{i-o}) is also evident (Fig. 2a). Paired oocytes expressing a Cx26P87L mutant, in which the probable distortion caused by the proline in M2 has presumably been relaxed, yielded no detectable intercellular conductance ($n=21$). But pairing a mutant-expressing oocyte with one expressing Cx26 wild type (designated Cx26/Cx26P87L) did produce intercellular coupling, although with sixfold lower conductance than wild-type pairs. Cx26/Cx26P87L pairs showed unique voltage gating properties (Fig. 2b), where $g_{j(ss)}$ increases with increasing positive V_j (positive defined henceforth with respect to the mutant-expressing oocyte), although this increase could not be fitted to a Boltzmann equation (see Fig. 2 legend). In most of the limited cases tested, gap junction hemichannels, including Cx26 (ref. 18), have been observed to close in response to V_j s positive at their cytoplasmic ends. It would then follow

* Present address: Department of Biochemistry, Nanjing University, Nanjing, People's Republic of China.

† To whom correspondence should be addressed.

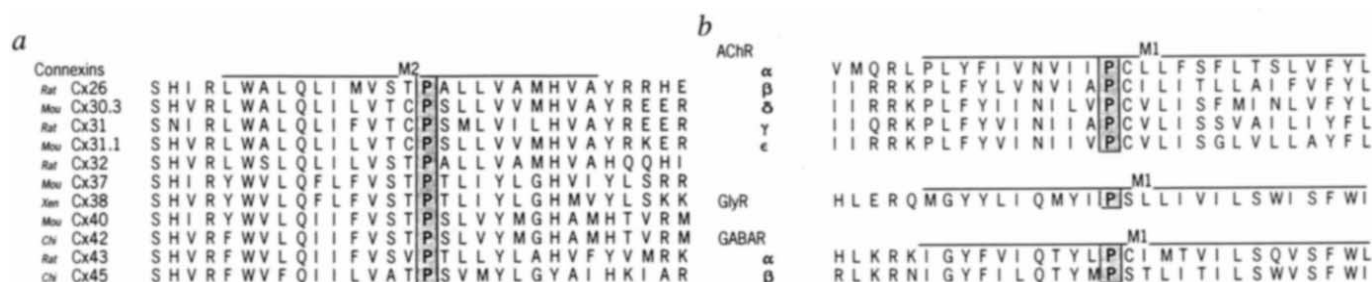
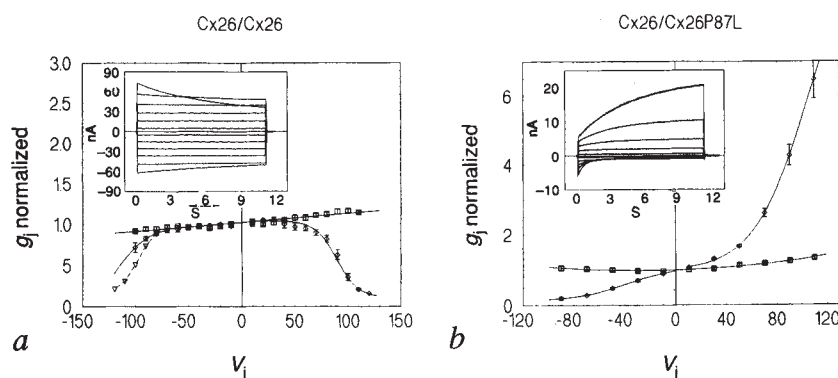


FIG. 1 Conserved proline residues are present in the second transmembrane-spanning segment (M2) of connexins and M1 of the ligand-gated channel superfamily. **a**, Sequence alignment of M2 from all members of the connexin family cloned so far²¹ shows a conserved proline in the middle of this helix. **b**, Alignment of the M1 regions of ligand-gated channels (mouse acetylcholine receptor subunits²², rat glycine receptor subunit²³, bovine GABA receptor subunits²⁴) with respect to

conserved proline residues. M1 of ligand-gated channels is analogous to M2 in connexins in that M1 immediately precedes the probable pore-forming region of ligand-gated channels and is oriented with its N terminus outside the cell. The subfamily of glutamate receptors (not shown here) are the least conserved of the ligand-gated channel superfamily with respect to sequence, and lack this consensus proline.

FIG. 2 Comparison of Cx26/Cx26 homotypic and Cx26/Cx26P87L heterotypic channel gating properties. Voltage sensitivity was determined by the change in junctional current in response to V_j which is applied as voltage steps in 20 mV increments ranging from +110 mV to -90 mV for Cx26 (**a**, inset) and +90 mV to -90 mV for Cx26/Cx26P87L (**b**, inset), from a holding potential of -45 mV. Pulses of 10 s are shown, but 30-s pulses were used to generate data for V_j versus g_j plots. Current traces were fitted with single exponential equations to determine time constants of decay ($\tau_c = 10.4$ s for Cx26/Cx26 ($V_j = +90$ mV) and 0.6 s for Cx26/Cx26P87L ($V_j = -90$ mV)) and activation ($\tau_o = 4.7$ s for Cx26/Cx26P87L ($V_j = +90$ mV)). Junctional conductance was measured within 10 ms after the onset of a voltage step ($g_{j(\text{init})}$) and at the end of a 30-s step ($g_{j(\text{ss})}$). Normalized values of $g_{j(\text{init})}$ (with respect to $g_{j(\text{init})}$ at 10 mV, $\square$) and $g_{j(\text{ss})}$ (with respect to $g_{j(\text{init})}$ at the corresponding V_j , $\diamond$) were plotted against V_j for Cx26/Cx26 (**a**) and Cx26/Cx26P87L (**b**) channels and fitted with Boltzmann curves where possible. V_j polarity is either defined as positive for depolarizing voltage steps in (**a**) or with respect to relative depolarizations of the Cx26P87L expressing oocytes in (**b**). (Note that because of the different ways that V_j is defined for the symmetrical Cx26/Cx26 and asymmetrical Cx26/Cx26P87L channels, the $g_{j(\text{init})}$ data are not directly comparable.) For Cx26/Cx26 channels, a holding potential of -20 mV was used to allow accurate determination of junctional conductance for $V_j < -100$ mV (V in **a**). Cx26/Cx26P87L oocyte pairs display ~sixfold lower average conductance (0.3 ± 0.1 μ S, $n = 24$) than Cx26/Cx26 pairs (1.9 ± 0.5 μ S, $n = 18$). All conductances quoted represent means $\pm$ s.e. The steady-state conductance increase in response to positive V_j for the Cx26/Cx26P87L channels could not accurately be fitted to a Boltzmann equation of the form $g_j = (g_{j(\text{max})} - g_{j(\text{min})}) / (1 + e^{(A(V_j - V_0))}) + g_{j(\text{min})}$ because an estimate of $g_{j(\text{max})}$ was not possible within the range of V_j studied ($\leq +120$ mV). Boltzmann constants describing the steady-state decay for Cx26/Cx26P87L channels at negative V_j ($A = 0.05$ and $V_0 = -39$ mV) differ significantly from



those describing the response of Cx26/Cx26 channels ($A = 0.1$ and $V_0 = +89$ mV).

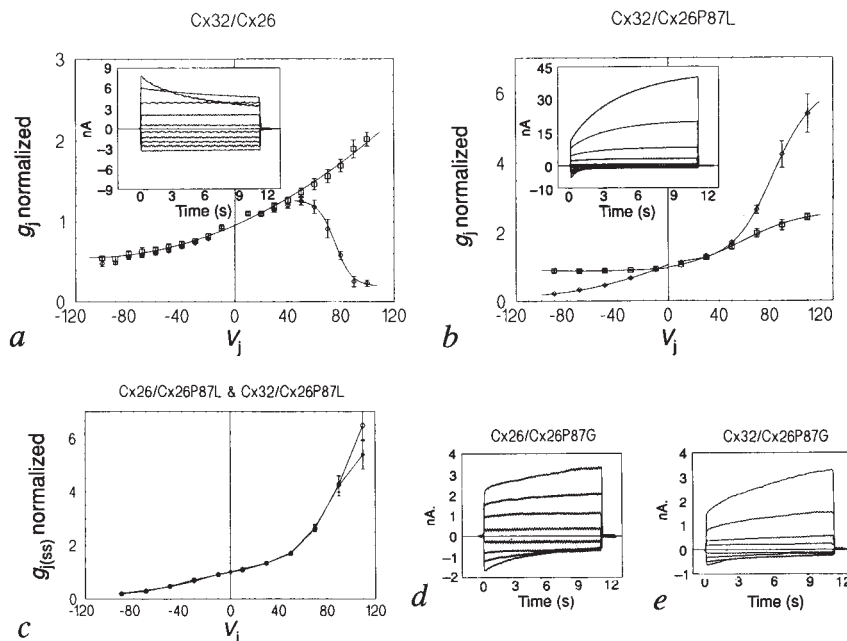
METHODS. The coding region of Cx26, with attached 5' and 3' *Xenopus* β -globin untranslated regions (described in refs 15, 25), was subcloned into the pGEM-7z vector. The Cx26P87L, as well as the G, A, I and V mutants described in Fig. 3 and the text, were prepared by the Kunkel method of point mutagenesis²⁶. The coding regions were fully sequenced in each case. cRNA transcripts prepared from the mutant clones by SP6 RNA polymerase (Promega, WI) were co-injected into stage V or VI *Xenopus* oocytes together with a 27-base antisense oligonucleotide complementary to nucleotides 327 to 353 of the *Xenopus* Cx38 coding region in order to eliminate contributions from endogenous coupling¹⁵. Injected oocytes were incubated for 24–48 h before stripping of their vitelline envelopes and manual pairing of vegetal poles. After a further 16–24 h, recordings were made using a dual electrode voltage clamp (Physiologic Instruments, CA) in each oocyte. Current and voltage electrodes with resistances of ~ 1 M Ω were filled with a 2 M KCl solution. Current traces were recorded on VCR tape, digitized by an Instrutech VR-100 A-D converter, and analysed on an IBM compatible 386 computer using Easy Plot (Spiral Software, MA).

that the conductance increase seen here is associated with the mutant hemichannel. Applied V_j of opposite polarity, which should affect the wild-type hemichannels, caused a reduction in $g_{j(\text{ss})}$, but with characteristics significantly different from those of wild-type channels (see Fig. 2 legend).

This was further investigated by pairing a mutant-expressing oocyte with one expressing Cx32, which appears to gate with the opposite polarity to Cx26 (the hemichannels close for V_j s relatively negative at their cytoplasmic ends¹⁸). Thus, heterotypic junctions between Cx32 and Cx26 only show steady-state gating when the Cx26 oocyte is relatively positive (and the Cx32 oocyte relatively negative) (Fig. 3a, and ref. 15). In contrast, Cx32/Cx26P87L heterotypic channels, which again have lower conductance than their wild-type counterpart (Fig. 3 legend), displayed a steady-state response to V_j that is virtually

indistinguishable from that of Cx26/Cx26P87L channels (Fig. 3c). This was also true of the time constants of both activation and inactivation (Fig. 3 legend). The only manifestation of Cx32 in Cx32/Cx26P87L pairings was the rectification in the plot of initial conductance ($g_{j(\text{init})}$) reminiscent of wild-type heterotypic channels (Fig. 3b). In light of the opposite gating polarities of Cx32 and Cx26, the superimposition of the steady-state responses of both Cx26/Cx26P87L and Cx32/Cx26P87L hybrid channels indicates that the entire response to V_j s of both polarities resulted from the mutant hemichannel. Our failure to observe the pure mutant channel is likely to result from the reduced conductance seen in wild-type/mutant pairings (6–20-fold below corresponding wild-type pairs) which might be predicted to be cumulative in the case of mutant channels. This would reduce conductances to the detection limit of this system (~ 0.002 μ S).

FIG. 3 Cx32/Cx26P87L and Cx26/ and Cx32/Cx26P87G channel formation. Cx32/Cx26P87L channels (mean $g_j = 0.21 \pm 0.73 \mu\text{S}$, $n = 23$) display ~20-fold lower conductance than wild-type Cx32/Cx26 channels (mean $g_j = 4.2 \pm 1.02 \mu\text{S}$, $n = 25$). In the V_j versus g_j plots in a and b, where the Cx26 expressing oocyte is defined as positive, the initial conductance of Cx32/Cx26P87L channels (b) display rectifying properties similar, although not identical, to those of Cx32/Cx26 channels (a). In contrast, the steady-state response of Cx32/Cx26P87L channels to V_j is greatly modified compared to the Cx32/Cx26 channels. Indeed, the voltage dependence of their conductance is not statistically different ($P > 0.2$, two-sided t -test) from the normalized response of Cx26/Cx26P87L channels, as evident in the superimposed graphs in c. Time constants for Cx32/Cx26P87L channel current decay (negative V_j) and increase (positive V_j) are: $\tau_c = 0.6 \text{ s}$ ($V_j = -90 \text{ mV}$) and $\tau_o = 4.2 \text{ s}$ ($V_j = +90 \text{ mV}$), again virtually identical to those of Cx26/Cx26P87L channels ($\tau_c = 0.6 \text{ sec}$; $\tau_o = 4.7 \text{ s}$ at the same V_j s). Current traces from oocytes expressing Cx26 (d) and Cx32 (e) paired with Cx26P87G injected oocytes have similar profiles to those for Cx26/ and Cx32/Cx26P87L pairs. But the mean conductance ($0.019 \pm 0.004 \mu\text{S}$, ($n = 7$) and $0.013 \pm 0.002 \mu\text{S}$, ($n = 10$) for Cx26P87G paired with Cx26 and Cx32, respectively) and the degree of current increase or decrease at positive or negative V_j s, respectively, is less in the Cx26P87G than the Cx26P87L mutant channels.



Even the increase in g_j with positive V_j associated with mutant hemichannels would be negated by the concomitant closure of the apposing mutant hemichannel.

To test the influence of the substituted residue, we also replaced Pro 87 with several amino acids in addition to leucine. Glycine differs from leucine in its molecular volume, and in its ability to take on backbone conformations similar to those of proline, if so constrained by the packing of surrounding structures. Nonetheless, both Cx26/Cx26P87G and Cx32/Cx26P87G pairings (Fig. 3d, e) produced qualitatively similar V_j responses to the analogous hybrids with Cx26P87L (compare inserts in Figs 2b and 3b), although the average g_j and the magnitude and rates of current activation and inactivation were reduced (Fig. 3). Substitution of several other residues (Ile, Ala and Val) for Pro 87 rarely produced a detectable g_j except in 1 out of 12 Cx32/Cx26P87A pairings. In this case, the conductance observed had a similar voltage gating profile to Cx32/Cx26P87G channels. Thus, the modified voltage gating properties of these mutants seem to be due to the removal of the structural conformation of M2 imposed by Pro 87, although the specific amino acid substituted does influence the ability to form functional (open) channels.

In wild-type gap junctions, increases in V_j cause a decrease in P_o (open probability) of the channels¹⁹, leading to a reduction in g_j . Although single channels could not be detected in the oocyte system, the results presented here are consistent with a mutant hemichannel for which P_o now increases with increasing

V_j (positive on the cytoplasmic side), although a decrease of relatively negative V_j s is also evident. The mutant gating properties could also be explained by voltage-dependent changes in the single channel conductance, analogous to what has been reported for Cx37 (ref. 20), or through recruitment of new channels over the time course of the voltage step. However, the magnitude and time course of these gating properties make these mechanisms less likely.

Although changes in Pro 87 may have far-reaching conformational consequences, many aspects of channel function are relatively unaffected. The channels assemble, dock with wild-type connexons, respond over a voltage range similar to wild-type channels, and close in response to lowered pH (the conductance of wild-type and wild-type/mutant channels decreased at a similar rate and to a similar degree in response to acetate-induced shifts in extracellular pH from 7.2 to 6.3; not shown). Thus, the distortion introduced into M2 by Pro 87 appears to play a role specific to voltage gating, perhaps by transducing movements of a 'voltage sensor' to movements of a channel gating element, or by acting as a 'scaffolding' about which specific gating elements are arranged. Substitution of a non-distorted M2 region in the mutant could lead to an increased rather than a decreased P_o in response to the same voltage change. With regard to other channels, it is notable that proline is also found in M1 of ligand-gated channels (Fig. 1). Thus, this residue may be a general motif used by different channels in their own context of gating. □

Received 27 May; accepted 27 August 1993.

1. Spray, D. C. & Bennett, M. V. A. *Rev. Physiol.* **47**, 281–303 (1985).
2. Hennemann, H. et al. *J. Cell Biol.* **117**, 1299–1310 (1992).
3. Milks, L. C. et al. *EMBO J.* **7**, 2967–2975 (1988).
4. Unwin, P. N. T. & Ennis, P. D. *Nature* **307**, 609–613 (1984).
5. Guy, R. H. & Conti, F. *Trends Neurosci.* **13**, 201–206 (1990).
6. Yun, R. H., Anderson, A. & Hermans, J. *Proteins Struct. Funct. Genet.* **10**, 219–228 (1991).
7. Barlow, D. J. & Thornton, J. M. *J. molec. Biol.* **201**, 601–619 (1988).
8. Deisenhofer, J. et al. *Nature* **318**, 618–624 (1985).
9. Henderson, R. et al. *J. molec. Biol.* **213**, 899–929 (1985).
10. Schulz, G. & Schirmer, R. H. *Principles of Protein Structure* (Springer, New York, 1979).
11. Brandl, C. J. & Deber, C. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 917–921, (1986).
12. Lolkema, J. S., Puttner, I. B. & Kaback, H. R. *Biochemistry* **27**, 8307–8310 (1988).
13. Mogi, T., Stern, L. J., Chao, B. H. & Khorana, H. G. *J. biol. Chem.* **264**, 14192–14196 (1989).
14. Fimmel, A. L. et al. *Biochem. J.* **213**, 451–458 (1983).

15. Barrio, L. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8410–8414 (1991).
16. Bennett, M. V. L. et al. *Gap Junctions* (eds Hertzberg, E. L. & Johnson, R. G.) 287–304 (Liss, New York, 1988).
17. Werner, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5380–5384 (1989).
18. Verselis, V. K. et al. *Neurosci. Abstr.* 22nd annual meeting, Anaheim CA, Oct. 25–30, Abstr. 275.10 (1992).
19. Veenstra, R. D. *Am. J. Physiol.* **258**, C662–C672 (1990).
20. Reed, K. E. et al. *J. clin. Invest.* **91**, 997–1004 (1993).
21. Haefliger, J. A. et al. *J. biol. Chem.* **267**, 2057–2064 (1992).
22. Connolly, L. G. *Comp. Biochem. Physiol.* **93A**, 221–231 (1989).
23. Grenningloh, G. et al. *Nature* **338**, 215–220 (1987).
24. Schofield, P. R. et al. *Nature* **328**, 221–227 (1987).
25. Willecke, K. et al. *J. Cell Biol.* **114**, 1049–1057 (1991).
26. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).

ACKNOWLEDGEMENTS. This work was supported by the NIH and by a PEW scholarship in the Biomedical Sciences awarded to B.J.N.

The protein IsK is a dual activator of K⁺ and Cl⁻ channels

Bernard Attali, Eric Guillemare, Florian Lesage, Eric Honoré, Georges Romey, Michel Lazdunski* & Jacques Barhanin

Institut de Pharmacologie Moléculaire et Cellulaire, CNRS, 660 route des Lucioles, Sophia Antipolis, 06560 Valbonne, France

* To whom correspondence should be addressed

THE protein IsK (*M_r* 14,500) is present in epithelial cells¹, heart^{2,3}, uterus⁴ and lymphocytes⁵ and induces slowly activating K⁺ currents when expressed in *Xenopus* oocytes¹. The finding that mutations of its single transmembrane segment altered channel gating⁶ or selectivity⁷ has suggested that IsK is a channel-forming protein. But IsK does not exhibit the K⁺ channel hallmarks⁸ (a conserved K⁺ selective pore (H5) flanked by either six^{9–11} or two^{12,13} membrane-spanning regions). Here we report that IsK expression in *Xenopus* oocytes also induces a Cl⁻ selective current very similar to the Cl⁻ current produced by phospholemman expression¹⁴ and with biophysical, pharmacological and regulation characteristics very different from those of the IsK-induced K⁺ channel activity. IsK mutagenesis identifies amino- and carboxy-terminal domains as critical for the induction of Cl⁻ and K⁺ channel activities, respectively. Our data lead to a model in which the IsK protein (now called IsK, Cl) acts as a potent activator of endogenous and otherwise silent K⁺ or Cl⁻ channels.

When the human IsK coding sequence was inserted into 3' and 5' untranslated regions of a *Xenopus* β -globin gene¹⁵, the resulting complementary RNA induced the expression of the expected slowly activating K⁺ currents upon membrane depolarization^{5,16} at low cRNA concentrations (threshold at 0.03 ng μ l⁻¹). In addition, an unexpected hyperpolarization-activated inward current was observed at high cRNA concentrations (-0.5 ± 0.3 μ A at -130 mV with 1 μ g μ l⁻¹; $n = 30$) (Fig. 1a). Its activation kinetics were very slow, with a threshold potential of -80 mV. A shift of the tail current reversal potential of 15 mV for a 106.5 to 55.6 mM change of the external Cl⁻ concentration indicated that this current was selective for Cl⁻ ions. Remarkably, the inward current was very similar to that induced by

expression in *Xenopus* oocytes of the recently cloned phospholemman¹⁴ (Fig. 1b, c). Moreover, equivalent amounts of IsK and phospholemman cRNA were required to generate this current (Fig. 1d). Phospholemman was, however, unable to induce depolarization-activated K⁺ currents.

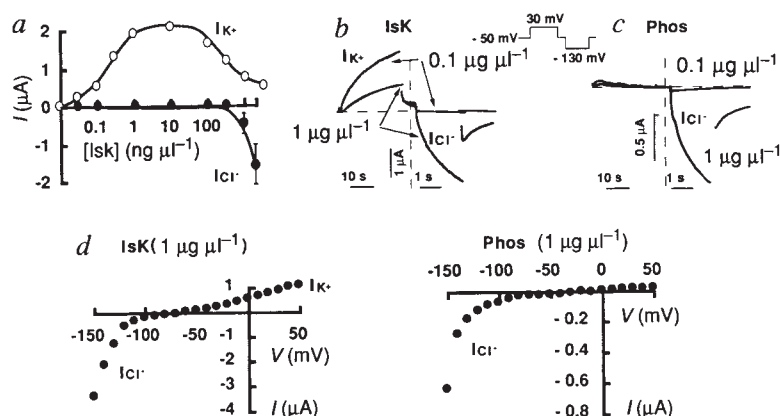
Chloride currents, induced either by IsK or phospholemman, had similar pharmacological properties distinct from those of the K⁺ current. DIDS (4,4'-diisothiocyanostilbene-2-2'-disulphonic acid) potentially inhibited both IsK and phospholemman-associated Cl⁻ currents but did not affect the K⁺ current (Fig. 2a). Barium was a potent blocker of both IsK- and phospholemman-induced Cl⁻ currents (50% inhibitory concentration, IC₅₀, was 0.4 mM and 0.3 mM, respectively) and a weaker blocker of the K⁺ current (IC₅₀ 3 mM). Clofilium, the most potent inhibitor of the IsK-associated K⁺ current^{2,3,5}, is an activator of both the IsK- (Fig. 2a) and phospholemman-induced Cl⁻ channel activities.

Increasing the osmolarity¹⁷ of the standard medium by 40 mOsm using saccharose or sorbitol decreased K⁺ channel activity by $69 \pm 3\%$ ($n = 4$). The same treatment was without effect on Cl⁻ channel activity obtained from either IsK or phospholemman.

IsK-associated K⁺ and Cl⁻ currents displayed differential modes of regulation. The protein kinase C (PKC) activator phorbol-12-myristate-13-acetate (PMA) decreased the K⁺ current^{18,19} but it strongly increased the IsK-associated Cl⁻ current (Fig. 2b). But PMA did not affect phospholemman-induced Cl⁻ current. By contrast, K⁺ currents were enhanced by raising internal Ca²⁺ (refs 2, 18; Fig. 2c), whereas Cl⁻ currents elicited by either IsK or phospholemman were increased by lowering external Ca²⁺ (Fig. 2c).

The capacity of IsK to produce either a K⁺ current or both K⁺ and Cl⁻ currents (depending on the cRNA concentration), the drastic differences in pharmacological and regulation properties and osmotic sensitivities of IsK-induced K⁺ and Cl⁻ currents, the difficulty of conceiving a pore structure with both a K⁺ > Rb⁺ > NH₄⁺ > Cs⁺ selectivity⁷ and a Cl⁻ selectivity, all strongly suggest that IsK is actually not a channel protein itself but rather an activator of endogenous and dormant K⁺ and Cl⁻ channels. This view is strengthened by the fact that IsK and phospholemman, two small proteins without sequence homology, give rise to the same Cl⁻ channel activity, and also by the failure to express IsK as a K⁺ channel after lipid bilayer reconstitution²⁰. A possible role of IsK (or phospholemman) in the induction of ion-channel gene transcription is unlikely because actinomycin D treatment of oocytes had no effect on

FIG. 1 Functional expression of IsK and phospholemman (Phos) in *Xenopus* oocytes. a, Dose-response relationship between current amplitudes and IsK cRNA concentrations. K⁺ and Cl⁻ currents were measured at +30 mV and -130 mV, respectively (holding potential: -50 mV). For high IsK cRNA quantities, a shift of the oocytes resting potential was consistently observed (from -59 ± 3 mV at 1 μ g μ l⁻¹ to -25 ± 2 mV at 3 μ g μ l⁻¹; $n = 20$). This shift could be due to the presence in the oocyte membrane of a small fraction of Cl⁻ channels opened at the resting potential. b, c, Superimposed current traces from oocytes injected with 0.1 and 1 μ g μ l⁻¹ IsK and phospholemman cRNA. Pulse-command protocol (inset): a 30-s depolarizing pulse is applied from the holding potential of -50 mV to +30 mV, then a 1-s step of repolarization to -50 mV, then a 3-s hyperpolarizing pulse to -130 mV. d, Current-potential relationships of IsK (1 μ g μ l⁻¹) and phospholemman (1 μ g μ l⁻¹)-induced currents. Oocytes were injected with IsK cRNA or phospholemman cRNA. IsK induced an outward K⁺ current on depolarization and an inward Cl⁻ current on hyperpolarization ($n = 5$). Phospholemman only induced the inward hyperpolarization-activated Cl⁻ current. Some batches (2 out of 25) of uninjected oocytes had hyperpolarization-activated Cl⁻ currents, suggesting that IsK and phospholemman may modulate an endogenous oocyte Cl⁻ channel.



METHODS. The IsK⁵ and the phospholemman²¹ cDNAs were obtained by polymerase chain reaction on human T lymphocyte and dog heart cDNAs, respectively, and inserted into the pBTG vector¹⁵. Methods for cRNA injection and electrophysiology have been described elsewhere²².

the expression of IsK-associated K^+ and Cl^- currents.

Mutagenesis was done to identify the structural elements involved in the activation of K^+ and Cl^- currents (Fig. 3). Replacing Ser 68 by Thr (S68T mutant) resulted in a complete loss of expression of the K^+ current with no alteration of the Cl^- channel activity (Fig. 3a, b). Truncation of the IsK C terminus (Tr80 mutant; Fig. 3a) resulted in a complete loss of K^+ channel activity with a striking increase of the Cl^- current amplitude (Fig. 3b). Conversely, the deletion of 28 residues at the IsK aminoterminal ($\Delta 11,38$ mutant) abolished the expression of the Cl^- current and doubled the amplitude of K^+ currents (Fig. 3b). The cRNA concentration dependency for the expression of K^+

and/or Cl^- currents by these mutants was unchanged as compared to wild-type IsK (not shown).

The K^+ current produced by a fixed concentration of the exclusive K^+ current inducer $\Delta 11,38$ mutant (Fig. 3c) and by the native IsK (not shown) were dose-dependently inhibited by the expression of S68T and Tr80 mutants. The cRNA concentrations of S68T and Tr80 that produced inhibition of K^+ channel

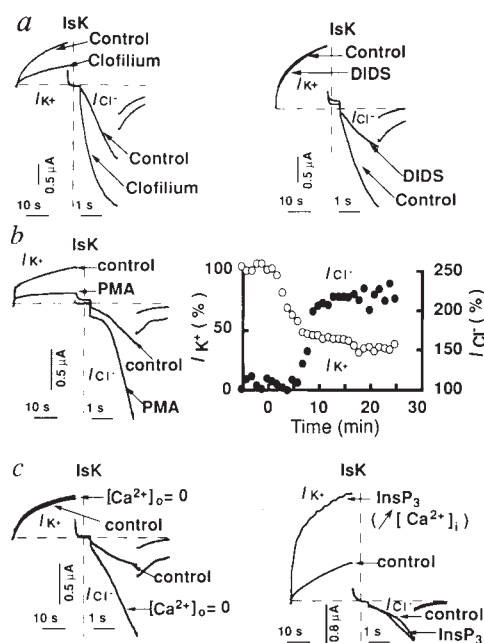


FIG. 2 Pharmacological and regulation properties of K^+ and Cl^- currents induced by IsK ($1 \mu g \mu l^{-1}$). a, Effects of clofilium ($100 \mu M$) and DIDS ($1 mM$). External SITS (4-acetamido-4'-isothiocyanostilbene-2,2'-disulphonic acid) ($1 mM$), 9-anthracene carboxylic acid ($2 mM$) and IAA 94/95 (5-nitro-2-(3-phenylpropylamino)-benzoic acid) ($2 mM$) and IAA 94/95 (indanyloxyacetic acid) ($1 \mu M$) were ineffective (not shown). b left, Effects of PMA ($30 nM$); right, kinetics of PMA action ($n=14$). Time 0 is the start of PMA perfusion. c left, Effects of extracellular Ca^{2+} : current traces were recorded in the presence of $1.8 mM CaCl_2$, $2 mM MgCl_2$ (control) or $0 mM CaCl_2$, $3.8 mM MgCl_2$ ($n=10$); right, effects of elevation of intracellular Ca^{2+} by injection of inositol trisphosphate ($InsP_3$) (final concentration: $1 \mu M$; $n=5$). Same current command protocol as for Fig. 1.

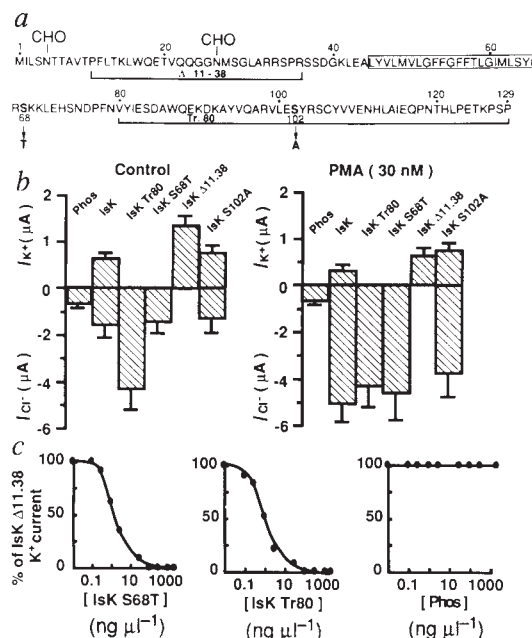


FIG. 3 Effects of mutations on IsK-induced channel activities and competition experiments. a, Point mutations, deletion and truncation are indicated below the amino-acid sequence of the human IsK protein. The putative transmembrane segment is boxed. Note that the N-terminal deletion $\Delta 11,38$ eliminates one putative N-glycosylation site and several positively charged residues lying upstream of the transmembrane domain. b, Mean amplitudes of K^+ and Cl^- currents generated by the expression of the different IsK mutants ($n=10$). Right, the mean current amplitudes of the same oocytes after a 30-min PMA treatment. c, Competition experiments were done by coinjection of $\Delta 11,38$ cRNA with S68T, Tr80 and phospholemman cRNA, respectively ($n=7$). The $\Delta 11,38$ cRNA concentration injected was $15 ng \mu l^{-1}$ and was saturating for the expression of K^+ current. Holding potential: $-50 mV$; pulse potentials: $+30 mV$ and $-130 mV$ for K^+ and Cl^- currents, respectively.

METHODS. A 240-bp $SacI$ - $AccI$ DNA fragment of the human IsK was subcloned into the pBTG vector to construct the Tr80 truncation mutant. Site-directed and deletion-mutagenesis were done using oligonucleotide primers according to the manufacturer's protocol (Promega). All mutational changes were checked by sequence determination.

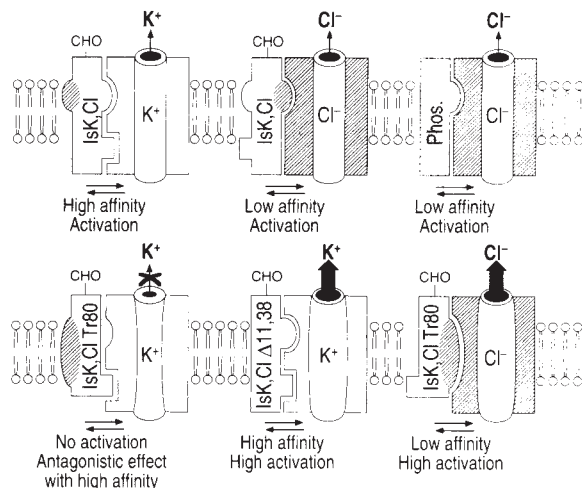


FIG. 4 Proposed model for K^+ and Cl^- channel activation by IsK, Cl and phospholemman (Phos). The high cRNA concentration necessary to induce Cl^- current suggests a low-affinity interaction between the Cl^- channel and the regulator entities. Conversely, the low cRNA concentrations necessary either to induce the K^+ current, or to inhibit it in competition experiments, suggests a high-affinity interaction between the K^+ channel and the regulator entities. The decrease of K^+ channel activity at high IsK cRNA concentrations (Fig. 1a) might suggest that (at least) two different types of oligomers could be formed between IsK and the endogenous K^+ channel, one being activatory and formed at low cRNA concentrations (depicted in the model) and the other being abortive and formed at high cRNA concentrations. The abortive complex would not be formed with the $\Delta 11,38$ mutant which does not display inhibitory behaviour at high cRNA concentrations (not shown).

expression were far below those necessary to induce Cl^- currents. Clearly, the single mutation S68T or the truncation of the C-terminal sequence have not abolished the capacity of the mutants to interact with a K^+ channel structure but have removed their capacity to act as activators. The mutants then behave as antagonists. In contrast, expression of phospholemman failed to prevent K^+ channel activity induced by $\Delta 11,38$ (Fig. 3c) or native IsK (not shown).

Replacement of Ser 102 (essential for PKC inhibition of K^+ channel currents¹⁹) by alanine suppressed PKC inhibition of the K^+ current but had no effect on the PKC enhancement of the Cl^- current (S102A; Fig. 3b). Therefore sites for PKC regulation of K^+ and Cl^- channel activities are clearly different, in agreement with a model in which IsK activates two distinct endogenous channels.

The model in Fig. 4 summarizes our views on the function of the IsK protein. The protein is now designated as IsK.Cl and is not considered as a K^+ forming pore protein. Rather, it behaves as a channel activator able to interact with endogenous K^+ and Cl^- channels but with a higher affinity for K^+ channels. The $\Delta 11,38$ mutant defines the IsK.Cl N-terminal region as a critical determinant for Cl^- channel activation. The Tr80 mutant defines the C-terminal region as crucial for the activation of the K^+ channel. Mutants (S68T and Tr80) that have lost their capacity to activate the K^+ channel can still bind to it but then behave as antagonists of IsK.Cl action. With this model, the only explanation for the modest changes of K^+ channel selectivity observed after an IsK F55T mutation⁷ is that IsK, acting as a regulator, has some contribution in defining the K^+ pore properties.

Phospholemman, a 72-amino-acid cardiac protein with a single transmembrane segment but no sequence homology with IsK.Cl, was considered to be a potential Cl^- channel protein¹⁴.

Our results strongly suggest that phospholemman is an activator of the same endogenous Cl^- channel that is revealed by IsK.Cl. We suggest that IsK.Cl and phospholemman are prototypic members of a new family of membrane proteins capable of having an all-or-none effect on cationic and/or anionic channel expression. □

Received 10 May; accepted 1 September 1993.

1. Takumi, T., Ohkubo, H. & Nakanishi, S. *Science* **242**, 1042–1045 (1988).
2. Honore, E. et al. *EMBO J.* **10**, 2805–2811 (1991).
3. Folander, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2975–2979 (1990).
4. Pragnell, M. et al. *Neuron* **4**, 807–812 (1990).
5. Attali, B. et al. *J. biol. Chem.* **267**, 8650–8657 (1992).
6. Takumi, T. et al. *J. biol. Chem.* **266**, 22192–22198 (1991).
7. Goldstein, S. A. N. & Miller, C. *Neuron* **7**, 403–408 (1991).
8. Aldrich, R. *Nature* **362**, 107–108 (1993).
9. Jan, L. Y. & Jan, Y. N. *A. Rev. Physiol.* **54**, 535–555 (1992).
10. Pongs, O. *Physiol. Rev.* **72**, S69–S88 (1992).
11. Rudy, B., Kentros, C. & Vega-Saenz De Miera, E. *Molec. cell. Neurosci.* **2**, 89–102 (1991).
12. Ho, K. et al. *Nature* **362**, 31–38 (1993).
13. Kubo, Y., Baldwin, T. J., Jan, Y. N. & Jan, L. Y. *Nature* **362**, 127–133 (1993).
14. Moorman, J. R., Palmer, C. J., John, J. E., Durieux, M. E. & Jones, L. R. *J. biol. Chem.* **267**, 14551–14554 (1992).
15. Kreig, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057–7070 (1984).
16. Murai, T., Kazikuka, A., Takumi, T., Ohkubo, H. & Nakanishi, S. *Biochem. biophys. Res. Commun.* **161**, 176–181 (1989).
17. Busch, A. E., Varnum, M., Adelman, J. P. & North, R. A. *Biochem. biophys. Res. Commun.* **184**, 804–810 (1992).
18. Honore, E., Attali, B., Lesage, F., Barhanin, J. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **184**, 1135–1141 (1992).
19. Busch, A. E., Varnum, M. D., North, R. A. & Adelman, J. P. *Science* **255**, 1705–1707 (1992).
20. Lesage, F. et al. *Receptors Channels* **1**, 143–152 (1993).
21. Palmer, C. J., Scott, B. T. & Jones, L. R. *J. biol. Chem.* **266**, 11126–11130 (1991).
22. Guillemare, E. et al. *Biochemistry* **31**, 12463–12468 (1992).

ACKNOWLEDGEMENTS. We thank L. Pradier (Rhône Poulenc Rorer) for the pBTG vector, K. Meisner for reading the manuscript, and M.-M. Larroque, C. Roulinat and F. Aguilera for technical assistance. This work was supported by the CNRS, the Association Française contre les Myopathies and the Association pour la Recherche sur le Cancer (ARC). B.A. was a recipient of an ARC fellowship.

Human xeroderma pigmentosum group D gene encodes a DNA helicase

Patrick Sung*, Véronique Bailly*†, Christine Weber‡, Larry H. Thompson‡, Louise Prakash*† & Satya Prakash*§

Department of Biology, University of Rochester, Rochester, New York 14627, USA

† Department of Biophysics, University of Rochester School of Medicine, Rochester, New York 14642, USA

‡ Lawrence Livermore National Laboratory, Biology and Biotechnology Research Program, Livermore, California 94550, USA

XERODERMA pigmentosum (XP), a genetically heterogeneous human disease, results from a defect in nucleotide excision repair of ultraviolet-damaged DNA. XP patients are extremely sensitive to sunlight and suffer from a high incidence of skin cancers. Cell fusion studies have identified seven XP complementation groups, A–G^{1–3}. Group D is of particular interest as mutations in this gene can also cause Cockayne's syndrome and trichothiodystrophy⁴. The *XPD* gene was initially named *ERCC2* (excision repair cross complementing) as it was cloned using human DNA to complement the ultraviolet sensitivity of a rodent cell line⁵. We have purified the XPD protein to near homogeneity and show that it possesses single-stranded DNA-dependent ATPase and DNA helicase activities. We tested whether *XPD* can substitute for its yeast counterpart *RAD3*, which is essential for excision repair and for cell

viability⁶. Expression of the *XPD* gene in *Saccharomyces cerevisiae* can complement the lethality defect of a mutation in the *RAD3* gene⁶, suggesting that *XPD* is an essential gene in humans.

The *XPD* gene was expressed in yeast and XPD protein was purified to a very high degree (~20,000 fold). Affinity-purified anti-XPD antibodies were used for immunoblotting to verify the expression of XPD protein in yeast and to monitor the elution of the protein from various chromatographic matrices during its purification (Fig. 1a). The size of the XPD protein produced in yeast agrees well with the M_r of 86.9K predicted from the nucleotide sequence⁵ and is identical to the size of XPD protein immunoprecipitated from human cell extract (Fig. 1b).

Purified XPD protein hydrolyses ATP to ADP and P_i only in the presence of single-stranded (ss) M13 or ss ΦX174 DNA, ATPase activity being much less with double-stranded (ds) M13 or pBR322 DNA (Table 1). XPD ATPase activity requires Mg^{2+} (Table 1), which can be substituted by Mn^{2+} but not by Ca^{2+} , Co^{2+} or Zn^{2+} (data not shown). Of all the other NTPs and dNTPs tested, only dATP was hydrolysed by the XPD protein, the efficiency and cofactor requirements being similar to ATP hydrolysis (data not shown). The ssDNA-dependent ATPase activity has a pH optimum of 6.5 (data not shown) and a K_m of 200 μM for ATP, as determined from a double-reciprocal Lineweaver–Burk plot using ssM13 DNA (60 μM nucleotide concentration). The ssDNA-dependent ATPase activity co-elutes with XPD protein during chromatography on a Mono-Q column and is absent in fractions derived from an extract of control yeast strain CMY135 lacking the *XPD* gene (Fig. 1c). Our results indicate that the ssDNA-dependent ATPase activity is intrinsic to XPD.

To determine whether XPD protein possesses a DNA helicase activity, it was incubated with ssM13 circular DNA to which a ³²P-labelled 74-nucleotide complementary fragment had been annealed. XPD catalyses the unwinding of the partial duplex in

* Present address: Sealy Center for Molecular Science, University of Texas Medical Branch, 6.104 Medical Research Bldg, J61, Galveston, Texas 77555, USA.

§ To whom correspondence should be addressed.

TABLE 1 XPD has ssDNA-dependent ATPase activity

Conditions	ATPase activity (%)
No DNA	0
Single-stranded DNA	
M13mp18	100
ΦX174	100
M13mp18 (–Mg ²⁺)	0
ΦX174 (–Mg ²⁺)	0
Duplex DNA	
M13mp18	5
pBR322	4
M13mp18 (–Mg ²⁺)	0
pBR322 (–Mg ²⁺)	0

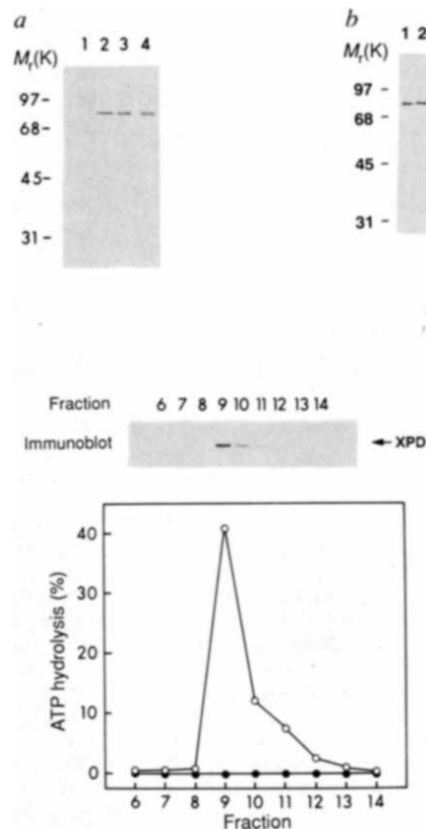
Purified XPD protein (20 ng) was incubated in 10 µl buffer R (20 mM potassium MES, pH 6.5, 100 µg ml⁻¹ BSA, 50 mM KCl, 5 mM MgCl₂, and 1 mM dithiothreitol) containing 0.5 mM [2,5,8-³H]ATP (10⁵ c.p.m. total; Amersham) and 200 ng DNA for 30 min at 37 °C. The duplex DNAs contained 80% supercoiled and 20% nicked-circular forms. ATP hydrolysis was quantitated by thin-layer chromatography using polyethyleneimine cellulose; 100% activity corresponds to the hydrolysis of 32% of the input ATP. Under these conditions, ATP hydrolysis is linear with time for at least 45 min.

FIG. 1 Purification of XPD and characterization of its ssDNA-dependent ATPase activity. **a**, Expression of XPD protein in yeast and its purification. The nitrocellulose blot of an 8% denaturing polyacrylamide gel (lanes 1 to 3) was probed with affinity-purified antibodies specific for XPD. Lane 1, whole-cell extract from strain CMY135 harbouring the plasmid pPM1 (2 µ, ADC1), which contains the ADC1 promoter but lacks the XPD gene; lane 2, extract from CMY135 harbouring plasmid pPM7 (2 µ, ADC1-XPD) carrying the XPD gene under the control of the ADC1 promoter; lane 3, 5 ng purified XPD protein. Lane 4, gel strip containing 100 ng purified XPD protein was stained with silver nitrate. **b**, Extract from normal human fibroblasts, cell line GM0637, was immunoprecipitated with protein A-agarose beads bearing anti-XPD antibodies. Bound proteins were eluted from the immunoprecipitate with 2% SDS and the eluate, equivalent to 10⁷ cells, was analysed by immunoblotting (lane 2) together with XPD protein purified from yeast (lane 1). **c**, ssDNA-dependent ATPase activity is intrinsic to XPD. Mono-Q column fractions 6 to 14 (0.8 ml each) from the final step of purification (see Methods) were each concentrated to ~100 µl using Centricon-30 (Amicon) and 0.8 µl of the concentrated fractions were assayed for ATP hydrolysis in pH 6.5 buffer as described in Table 1 legend. Open circles, concentrated Mono-Q fractions using extract from strain CMY135[pPM7] carrying the XPD gene; filled circles, concentrated Mono-Q fractions using extract from strain CMY135[pPM1], which lacks the XPD gene. ATPase activity closely parallels the amount of XPD as shown in the immunoblot.

METHODS. SDS-PAGE, immunoblotting, and staining of gels with silver nitrate were carried out as described⁸. Rabbit polyclonal antiserum was raised against a p-XPD hybrid protein expressed using the λP₁ promoter in *E. coli*. The insoluble polypeptide contains 406 amino acids, 353 of which are derived from XPD. Antibodies were purified by affinity chromatography. To express XPD protein in yeast, the cDNA encompassing the entire open reading frame of ERCC2(XPD)⁵ was placed under the control of the ADC1 promoter in the yeast 2 µ multicopy vector pPM1 (2 µ, ADC1 promoter, TRP1) to yield plasmid pPM7. Yeast strain CMY135 [pPM7] was precultured in complete synthetic medium lacking tryptophan and grown in YPD medium in 10-litre batches in fermenters at 30 °C to a density of 7 × 10⁷ cell ml⁻¹. All subsequent steps were carried out at 4 °C. Whole extract was prepared from 450 g yeast cells using a French press and, after high-speed centrifugation (100,000g for 120 min), the supernatant was treated with ammonium sulphate (0.21 g ml⁻¹). Protein precipitate was collected by centrifugation (20,000g, 15 min), dissolved in 300 ml buffer K (20 mM KH₂PO₄, pH 7.5, containing 10% glycerol and 1 mM dithiothreitol (DDT), and dialysed overnight, with one change, against a total of 8 litres of buffer K. The dialysate was clarified by centrifugation (20,000g, 15 min), passed through S-Sepharose (2.5 × 24 cm), applied onto Q-Sepharose (5 × 16 cm), which was washed with 600 ml 30 mM KCl in buffer K, and eluted with 300 mM KCl in buffer K; the protein peak (fraction 1; 200 ml) was treated with ammonium sulphate (0.25 g ml⁻¹). The ammonium sulphate protein precipitate was collected by centrifugation (20,000g, 30 min), dissolved in 80 ml buffer K, and dialysed against 2 litres buffer

a time-dependent (Fig. 2a) and protein-concentration-dependent fashion (Fig. 2b). The DNA helicase activity is maximal at pH 6.5 (data not shown). DNA unwinding shows a strict requirement for Mg²⁺ (data not shown), and for ATP (Fig. 2c; compare lanes 3 and 2), which can be substituted efficiently by dATP (Fig. 2c, lane 4) but not by CTP, GTP or UTP (Fig. 2c, lanes 6–8). ATP-γS, a non-hydrolysable ATP analogue, does not support DNA unwinding (Fig. 2c, lane 5), suggesting that ATP hydrolysis is necessary to drive the unwinding reaction.

All known DNA helicases unwind duplex DNA with a specific polarity owing to their unidirectional movement on ssDNA⁷. To determine the directionality of the XPD helicase, we constructed a linear substrate that consists of ssM13 DNA with terminal duplex regions of 18 base pairs(bp) and 34 bp; displacement of either the 18-nucleotide fragment or the 34-nucleotide fragment would indicate a 3' to 5' or a 5' to 3' direction of translocation of the helicase on the ssDNA, respectively (Fig. 2d). As positive controls, we included the *Escherichia coli* UvrD protein⁷ (3' to 5' directionality) and the yeast RAD3 protein⁸ (5' to 3' directionality) in our analysis (Fig. 2d; lanes 3 and 4). XPD displaces the 34-nucleotide fragment but not the 18-nucleotide fragment from the substrate (Fig. 2d, lanes 5–7), indicating that it translocates on bound ssDNA in the 5' to 3' direction.



K for 3 h to lower the ionic strength to 40 mM KCl. The dialysate was applied onto S-Sepharose (1.6 × 8 cm), which was washed with 60 ml 30 mM KCl in buffer K and developed with a 120-ml gradient of 50–225 mM KCl. The XPD peak (fraction 2, 15 ml), eluting at ~180 mM KCl, was identified by immunoblotting and dialysed against 1 litre buffer K for 2 h to lower the ionic strength to 40 mM KCl. The dialysate was chromatographed in a Mono-S column (HR5/5) with a 20-ml gradient of 50–250 mM KCl. The XPD peak (fraction 3, 0.8 ml), eluting at ~160 mM KCl, was size-fractionated in a Sephacryl-S100 column (1 × 45 cm) in buffer K containing 30 mM KCl. The S100 pool (fraction 4, 3 ml) was chromatographed on Mono-Q (HR5/5) using a 12-ml gradient of 30–200 mM KCl, and fractions of 0.8 ml were collected. The Mono-Q pool (fraction 5; 5 µg XPD) was concentrated using Centricon-30 (Amicon) and stored in small aliquots at –70 °C.

The *XPD* and *S. cerevisiae RAD3* encoded proteins are very homologous, having 52% identical residues⁵. In addition to its role in excision repair, *RAD3* is essential for cell viability⁶. In view of the high degree of evolutionary conservation, we determined whether XPD protein can provide these functions missing in *rad3* mutants. We introduced plasmid pPM7, which expresses the *XPD* gene, into the diploid yeast strain YR3-20 (*rad3-2/rad3Δ*); as controls, strain YR3-20 was transformed either with the plasmid pSCW367, which carries the *RAD3* gene fused to the *ADC1* promoter, or with the vector plasmid pPM1, which contains only the *ADC1* promoter. These diploid strains were sporulated and the genotypes of the spores in tetrads determined. As a result of the lethality of the *rad3Δ* mutation, YR3-20[pPM1] produced tetrads with only two viable spores (Fig. 3a, middle panel) which were sensitive to ultraviolet radiation because of the *rad3-2* mutation. A majority of tetrads from YR3-20[pSCW367] yielded four viable spores (Fig. 3a, top panel), indicating complementation of the inviability defect of *rad3Δ* spores by the *RAD3* gene. Interestingly, YR3-20[pPM7] produced many tetrads in which all four spores were viable; but two of the spores in the tetrad produced colonies of normal size and the other two spores produced smaller colonies (Fig. 3a, bottom panel). All four spores carried the plasmid pPM7. Genetic analysis confirmed that small colonies arose from spores that carry the *rad3Δ* mutation and larger colonies were derived from the *rad3-2* containing spores. These observations indicate that the *XPD* gene is able to substitute in the viability function of *RAD3* in yeast but that it is not as efficient as the yeast *RAD3* gene. In liquid YPD medium, we estimate the doubling time

of the haploid *rad3Δ*[pPM7] strain to be ~7 h, whereas the doubling time is 1.5 h for the isogenic *RAD3* strain.

Determination of the ultraviolet sensitivity of the *rad3Δ*[pPM7] strain indicated that *XPD* does not complement the excision repair defect of the *rad3Δ* mutation (data not shown), nor does it complement the ultraviolet sensitivity of either the incision-defective *rad3-2* mutant⁹ or the yeast *rad3 Arg-48* mutation¹⁰ (Fig. 3b), which confers a defect in excision repair but has no perceptible effect on viability. That *XPD* does not complement the excision repair defect of *rad3* mutants may indicate that XPD is unable to participate in productive interactions with the protein components of the excision repair machinery.

XPD complements the lethality of the yeast *rad3Δ* mutation, although it is much less effective than *RAD3*. Studies with a *rad3* mutant temperature-sensitive for growth have indicated a requirement for *RAD3* in RNA polymerase II-dependent transcription *in vivo*, which probably constitutes the essential role of the gene (our unpublished results). The ability of XPD to substitute for *RAD3* in its viability role suggests that domains controlling interactions between XPD and the transcriptional machinery have been evolutionarily conserved from yeast to humans.

In eukaryotes, transcription initiation by RNA polymerase II requires energy which can be provided by the hydrolysis of the β - γ bond of ATP or dATP^{11,12}. A DNA-dependent ATPase activity has been found in association with transcription factor b from yeast¹³, faction δ from rat¹⁴, and factor TFIIF in humans¹⁵. The *XPB(ERCC-3)* gene product copurifies with

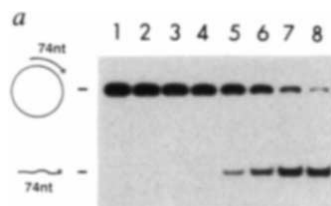
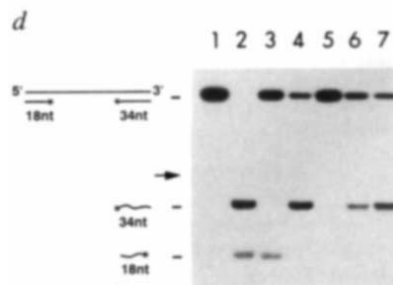
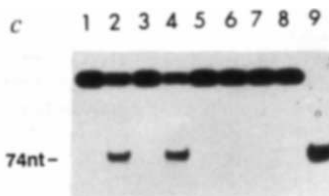
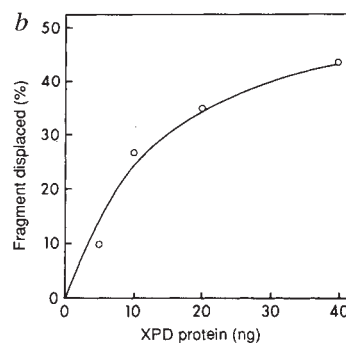


FIG. 2 XPD is a DNA helicase. a, Time-dependent unwinding of the 74-nucleotide (nt) partial duplex by the XPD protein. DNA substrate (lane 1) was incubated with 20 ng XPD protein for 2, 4, 6, 8, 10, 15 and 20 min (lanes 2 to 8, respectively). b, XPD concentration-dependent unwinding of DNA. The incubation time for this experiment was 10 min. c, Nucleotide requirement of XPD helicase. DNA substrate (lane 1) was incubated with 20 ng XPD protein for 10 min without any nucleotide (lane 3); with 1 mM of ATP (lane 2); dATP (lane 4); ATP- γ S (lane 5); CTP (lane 6); GTP (lane 7); and UTP (lane 8). In lane 9, DNA substrate was heat-denatured to release the hybridized ³²P-labelled fragment. d, XPD helicase has a 5' to 3' directionality. The linear substrate used for determining the direction of XPD translocation on ssDNA (lane 1) was heat-denatured (lane 2) or incubated for 10 min with 5, 10 or 20 ng XPD protein (lanes 5–7). The DNA substrate was also incubated for 10 min with 50 ng *E. coli* UvrD protein (lane 3) and 50 ng yeast *RAD3* protein (lane 4), which have 3' to 5' and 5' to 3' directionality, respectively. The structures of the circular and linear substrates are shown on the left in a and d respectively; asterisk denotes the ³²P label; arrow in d indicates the trace amount of circular substrate not cleaved by *Rsa*I.

METHODS. The 74-bp partial duplex was constructed by hybridizing a 69-nt-long fragment to viral ssM13 mp18 DNA and extending the 3' end of the hybridized fragment using dATP and [α -³²P]dCTP as described⁸. To prepare the directionality substrate used in d, a 41-base oligonucleotide complementary to positions 6,822–6,862 of viral ssM13 mp18 was labelled at the 5' end with [γ -³²P]ATP and polynucleotide kinase and hybridized to viral ssM13mp18 DNA. The 3' end of the hybridized oligonucleotide was extended with dATP and [α -³²P]dCTP using the *E. coli* DNA polymerase I (Klenow large fragment). The partial duplex was then treated with the restriction enzyme *Rsa*I to yield the linear substrate with terminal duplex regions of 34 bp and 18 bp. DNA helicase substrates were purified by gel filtration using a Sepharose-6B



column in 10 mM Tris-HCl, pH 7.0, 50 mM KCl and 0.1 mM EDTA. DNA unwinding by XPD was assayed in 10 μ l buffer R (Table 1 legend) containing 5 ng DNA helicase substrates and 1 mM ATP as described⁸.

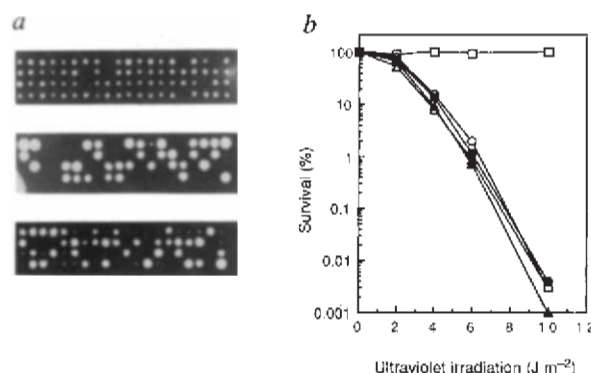


FIG. 3 XPD complements the inviability defect but not the DNA repair defect of *rad3* mutants. **a**, Tetrad analysis of strain YR3-20 (*rad3-2/rad3Δ*) with or without the XPD gene. Top, complementation of inviability of *rad3Δ* spores by the *RAD3* gene. Middle, only the two *rad3-2* spores survive, the *rad3Δ* spores are inviable. Bottom, XPD complements the inviability defect of *rad3Δ* spores. The slow-growing colonies are from the *rad3Δ* spores carrying the XPD gene on plasmid pPM7. Spores for the top panel were incubated for a shorter period than those shown in the middle and bottom panels. **b**, Ultraviolet sensitivity of *rad3* mutants, with or without the XPD gene. □, *RAD3* strain DBY747; ▲, *rad3-2* strain YR3-1 carrying the vector plasmid pPM1; Δ, *rad3-2* strain YR3-1 carrying the XPD plasmid pPM7; ●, *rad3 Arg-48* strain YR3-22 carrying the vector plasmid pPM1; ○, *rad3 Arg48* strain YR3-22 carrying the XPD plasmid pPM7.

TFIIH, and it has been suggested that the DNA helicase activity found in TFIIH is due to XPB¹⁵. Like *RAD3*, *RAD25*, the *S. cerevisiae* homologue of XPB, is required for excision repair and is essential for cell viability^{6,16,17}. Because of the similarity of *RAD3* and *RAD25*, they, as well as their respective human counterparts XPD and XPB, may function together in unwinding duplex DNA at the site of transcription initiation by RNA polymerase II. During excision repair, the XPD and XPB proteins may effect the localized unwinding of the damaged duplex DNA and promote repair synthesis after incision. □

Received 14 May; accepted 20 August 1993.

1. Cleaver, J. E. & Kraemer, K. H. in *The Metabolic Basis of Inherited Disease* (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 2949–2971 (McGraw-Hill, New York, 1989).
2. De Weerd-Kastelline, E. A., Keljzer, W. & Bootsma, D. *Nature* **238**, 80–83 (1972).
3. Vermeulen, W., Stefanini, M., Giliani, S., Hoejmackers, J. H. & Bootsma, D. *Mutat. Res.* **255**, 201–208 (1991).
4. Johnson, R. T. & Squires, S. *Mutat. Res.* **273**, 97–118 (1992).
5. Weber, C. A., Salazar, E. P., Stewart, S. A. & Thompson, L. H. *EMBO J.* **9**, 1437–1447 (1990).
6. Prakash, S., Sung, P. & Prakash, L. in *The Eukaryotic Nucleus* (eds Strauss, P. R. & Wilson, S. H.) 275–292 (Telford, Caldwell, New Jersey, 1990).
7. Matson, S. W., Kaiser-Rogers, K. A. & Rev. Biochem. **58**, 289–329 (1990).
8. Sung, P., Prakash, L., Matson, S. W. & Prakash, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8951–8955 (1987).
9. Wilcox, D. R. & Prakash, L. *J. Bact.* **148**, 618–623 (1981).
10. Sung, P., Higgins, D., Prakash, L. & Prakash, S. *EMBO J.* **7**, 3263–3269 (1988).
11. Bunick, D., Zandomeni, R., Ackerman, S. & Weinman, R. *Cell* **29**, 877–886 (1982).
12. Sawadogo, M. & Roeder, R. G. *J. Biol. Chem.* **259**, 5321–5326 (1984).
13. Feaver, W. J., Gileadi, O., Li, Y. & Komberg, R. D. *Cell* **67**, 1223–1230 (1991).
14. Conaway, R. C. & Conaway, J. W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7356–7360 (1989).
15. Schaeffer, L. et al. *Science* **260**, 58–63 (1993).
16. Gulyas, K. D. & Donahue, T. F. *Cell* **69**, 1031–1042 (1992).
17. Park, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11416–11420 (1992).

ACKNOWLEDGEMENTS. This work was supported by the US National Cancer Institute and the Department of Energy.

Phosphorylated CREB binds specifically to the nuclear protein CBP

John C. Chrivia[†], Roland P. S. Kwok^{*}, Ned Lamb[†], Masatoshi Hagihara[§], Marc R. Montminy[§] & Richard H. Goodman^{*}

^{*} Vollum Institute, Oregon Health Sciences University, Portland, Oregon 97201, USA

[†] Centre de Recherches de Biochimie Macromoléculaire, Centre National de la Recherche Scientifique-INSERM, 34033 Montpellier, France

[§] The Clayton Foundation Laboratories for Peptide Biology, Salk Institute, La Jolla, California 92037, USA

Cyclic AMP-regulated gene expression frequently involves a DNA element known as the cAMP-regulated enhancer (CRE)^{1–4}. Many transcription factors bind to this element, including the protein CREB^{5,6}, which is activated as a result of phosphorylation by protein kinase A⁷. This modification stimulates interaction with one or more of the general transcription factors or, alternatively, allows recruitment of a co-activator. Here we report that CREB phosphorylated by protein kinase A binds specifically to a nuclear protein of *M_r* 265K which we term CBP (for CREB-binding protein). Fusion of a heterologous DNA-binding domain to the amino terminus of CBP enables the chimaeric protein to function as a protein kinase A-regulated transcriptional activator. We propose that CBP may participate in cAMP-regulated gene expression by interacting with the activated phosphorylated form of CREB.

To identify proteins that interact with phosphorylated CREB, we screened a human thyroid λ gt11 library with recombinant

CREB labelled at the protein kinase A (PKA) phosphorylation site with [γ -³²P]ATP. One positive clone was identified among 4×10^5 plaques screened. This complementary DNA was used to screen a mouse brain cDNA library. The composite sequence of the predicted 265K mouse CBP contained several calmodulin kinase II phosphorylation sites, a single PKA phosphorylation site, two regions resembling C₄ zinc-fingers, and a carboxy-terminal glutamine-rich domain (Fig. 1a, b). Comparison with sequences in the EMBL and GENBANK databases did not reveal any proteins with extensive overall homology, but it did identify a region within the yeast coactivator ADA-2 (ref. 8) which was strongly related to the second zinc-finger of CBP (Fig. 1c). The subcellular localization of CBP was determined by indirect immunofluorescence using an affinity-purified anti-peptide antibody, which showed that CBP expression is largely restricted to the nucleus, although there was also a small amount of cytoplasmic staining (Fig. 2). Preabsorption of the antibody with synthetic peptide blocked all staining (data not shown).

To determine how phosphorylation influences CREB-CBP interactions, we used a sandwich binding assay. PKA-phosphorylated or non-phosphorylated CREB was incubated with a ³²P-labelled CRE oligonucleotide and used to probe a 58K N-terminal fragment of CBP blotted onto nitrocellulose filters (Fig. 3a). When non-phosphorylated CREB was used, no CBP was detected. In contrast the phosphorylated CREB complex clearly did detect the CBP fragment, indicating that binding occurs only if CREB is phosphorylated. To exclude nonspecific binding of radiolabelled oligonucleotide to CBP, we also tested labelled CRE alone as a probe. As expected, this probe bound to CREB but not CBP.

Co-immunoprecipitation was used to determine the specificity of this interaction. Casein kinase II phosphorylates CREB on three serine residues but this phosphorylation does not result in transcriptional activation (ref. 6, and M.R.M., unpublished observations). We therefore phosphorylated CREB with either PKA or casein kinase II, incubated it with a radiolabelled 78K

[†] Present address: St Louis University School of Medicine, Department of Pharmacological and Physiological Science, St Louis, Missouri 63104, USA.

2

[illegible]

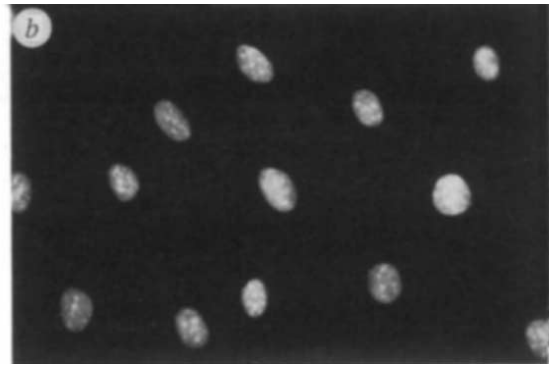
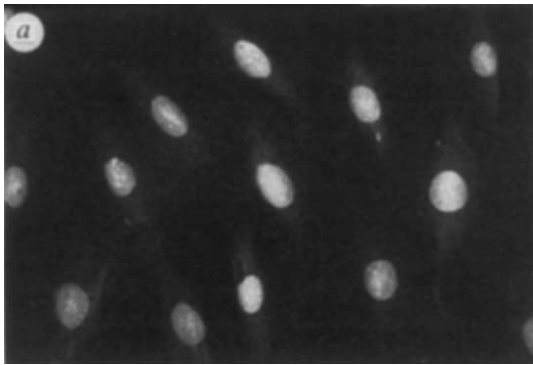


FIG. 2 Subcellular localization of CBP. REF-52 cells were fixed and incubated with anti-CBP antibody, Hoescht stain, or anti-CBP antibody preabsorbed with CBP peptide. Fluorescent micrographs are shown of cells stained for a, CBP or b, DNA. Staining patterns were similar for three other anti-CBP antibodies.

METHODS. Rabbit anti-CBP antibodies were raised against a peptide corresponding to the N-terminal 22 amino acids of CBP

(AENLLDGPPNPKRAKLSSPGFS)¹⁴ and purified by passing sera over a CBP-peptide affi-15 column (Biorad). Rat embryo fibroblast (REF-52) cells were cultured on glass coverslips, fixed in formalin, and extracted in acetone as described¹⁵. Cells were stained with affinity-purified antisera diluted 1/100 in PBS BSA for 60 min at 37 °C. Primary antisera were visualized with fluorescein-conjugated goat anti-rabbit serum and cells were stained for DNA, mounted, and photographed¹⁵.

CBP fragment, and immunoprecipitated the complexes with anti-CBP antibodies (Fig. 3b). Only the PKA-phosphorylated CREB co-precipitated with CBP, indicating that binding to CBP requires the specific phosphorylation of CREB at serine 133. CBP did not bind to other proteins phosphorylated by PKA, such as c-Fos or the R_{IIa} subunit of PKA (data not shown).

To identify the region of CREB involved in CBP binding, we expressed CREB residues 1–160 in bacteria as a glutathione-S-transferase (GST) fusion protein. This fusion protein lacks the basic-leucine zipper domain (Fig. 3c). The CREB1–160 peptide was purified, radiolabelled using [γ -³²P]ATP and PKA, and used to probe nitrocellulose filters containing electroblotted CBP. CREB1–160 detected immobilized CBP, indicating that the CREB activation domain was sufficient for binding. To define this domain more precisely, we examined a series of fusion proteins containing various activation-domain deletions. Transactivation properties of these deletion mutants had been determined⁹. CREBAQ1, which lacks the N-terminal 87 amino acids of CREB, still binds CBP. This deletion does not block PKA activation of CREB *in vivo*. CREBAQ2, which lacks amino acids 160–283, also retains the ability to bind CBP. The fusion protein GST-CREB101–255 recognized both immobilized and soluble CBP (Fig. 3c, d). Collectively, these results indicate that amino acids 101–160 within the CREB activator domain are sufficient for CBP binding.

Because deletions between amino acid residues 136–160 abolish the transactivation properties of CREB^{9,10}, we investigated whether these mutations also block CBP binding. Although CREB1–283 (Δ 140–160) can be phosphorylated by PKA, this transcriptionally inactive mutant does not bind immobilized or soluble CBP (Fig. 3c, d). A phosphopeptide representing only residues 128–141 of CREB does not bind CBP either, indicating that structural motifs outside the PKA phosphorylation site are required for CBP binding.

To test the functional properties of CBP, we constructed an expression vector containing the DNA-binding domain of the transcription factor GAL4 (amino acids 1–147) fused to the CBP

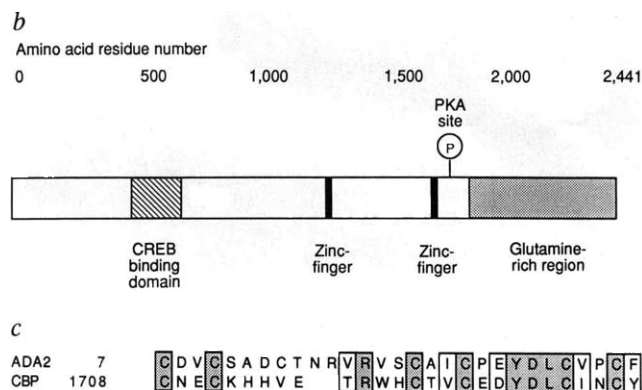


FIG. 1 Structure of CBP. a, Nucleotide sequence of the coding region of mouse CBP is shown, with the predicted amino-acid sequence in single-letter code. The differences between mouse and human CBPs (amino acids 1–1,267) are indicated: additional amino acids inserted in the mouse sequence that are not found in the human sequence (amino acids 808–810) are indicated by dashes; amino acids in the human CBP sequence that are deleted in the mouse protein are indicated by an asterisk, consensus calmodulin (CaM) kinase II sites by two asterisks, and the single consensus PKA site by three asterisks; the CREB-binding domain (amino acids 462–661) is underlined. Two putative C₄ zinc-fingers are indicated by the dashed double lines. b, The relative positions of the CREB-binding domain (determined by deletion analysis; data not shown), the putative zinc-fingers, the consensus PKA site and the glutamine-rich region in CBP are shown. c, Comparison of homologous regions in the yeast co-activator protein ADA-2 (amino acids 7–41) and CBP (amino acids 1,708–1,740). Shaded boxes indicate identical residues; unshaded boxes indicate conservative substitutions.

METHODS. A human Graves' disease thyroid λ gt11 cDNA library was plated at a density of 2×10^4 PFU per 150-mm dish and protein expression was induced by overlaying with nitrocellulose filters impregnated with isopropyl β -D-thiogalactopyranoside (IPTG)¹¹. Filters were removed, dried quickly, and preincubated for 1 h with blocking buffer (5% non-fat powdered milk, 1% BSA, 2.5% PVP-40, 0.1% Triton X-100 in PBS). Filters were incubated for 2 h at room temperature with 500 ng radiolabelled recombinant CREB in modified buffer A¹² and washed in the same buffer. Positive clones were purified and used to rescreen the thyroid library and a random-primed λ gt11 cDNA library prepared from mouse brain (Clontech). Isolated clones were sequenced by the dideoxy method¹³.

amino terminus. This vector activated a GAL-CAT reporter 10-fold (Fig. 4). Co-transfection of a plasmid encoding the catalytic subunit of PKA stimulated expression 45-fold, indicating that the transactivation potential of CBP could be regulated by PKA.

Two observations indicate that CBP does not function by recruiting endogenous CREB to the promoter. First, addition of exogenous CREB reduces, rather than stimulates, activity of the GAL-CAT reporter (Fig. 4). Second, GAL-CBP3, a vector containing the GAL4 binding domain fused to the CREB-

binding portion of CBP, was only minimally active when transfected either alone or with PKA. GAL-CBP3 was slightly more active in the presence of exogenous CREB and PKA, however, suggesting that some recruitment can occur if expression of exogenous CREB is high.

In summary, our data indicate that CBP has many of the features that would be expected of a CREB co-activator. Our findings would fit a model in which phosphorylation by PKA enables CREB to serve as a scaffold for CBP, which then activates the processes responsible for inducing expression of cAMP-

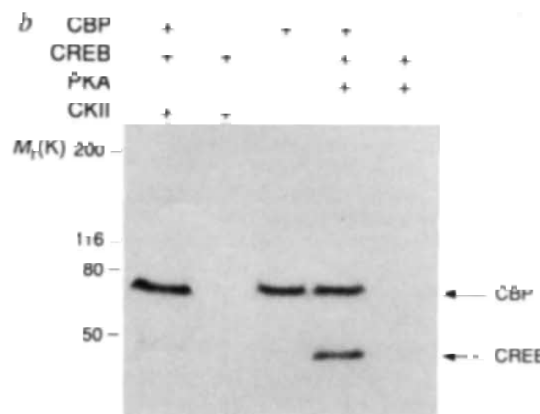
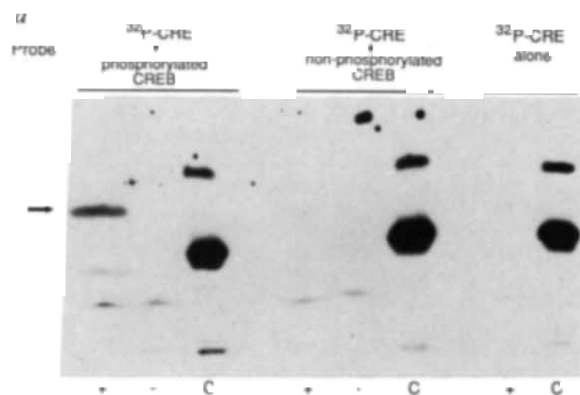
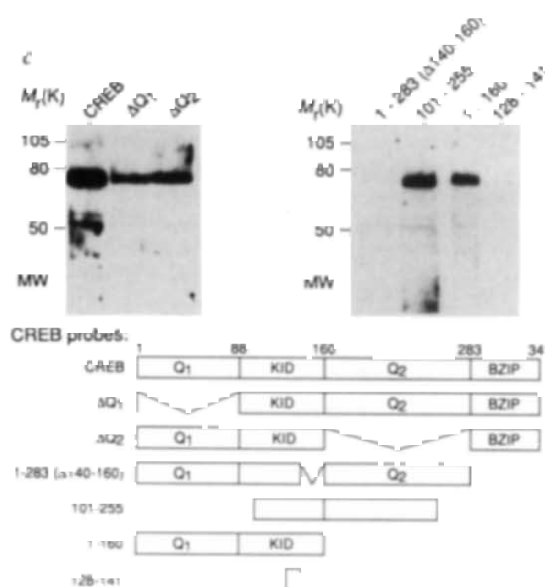


FIG. 3 CBP binds specifically to the activation domain of CREB phosphorylated by PKA. *a*, Complexes containing a radiolabelled CRE and either non-phosphorylated or PKA-phosphorylated CREB were used to probe western blots. 'Plus' lanes contain extracts from bacteria expressing a 58K CBP fragment (amino acids 117-661); 'minus' lanes contain extracts from control bacteria; C lanes show extracts from bacteria expressing CREB. The CBP band is indicated by an arrow. *b*, A 78K CBP fragment (amino acids 1-661), engineered to contain a PKA phosphorylation site at the C terminus was labelled with ^{32}P and mixed with ^{32}P CREB (labelled using either PKA or casein kinase (CK)II). Complexes were analysed by immunoprecipitation, SDS-PAGE and autoradiography. *c*, Bacterial extracts expressing a 78K N-terminal fragment of CBP were subjected to SDS-PAGE, electroblotted onto nitrocellulose, and probed with radiolabelled CREB proteins as indicated. *d*, GST-CREB fusion proteins were phosphorylated with PKA and immobilized on a glutathione column. Binding of the radiolabelled CBP fragment was detected only on the column containing GST CREB101-255, and only after phosphorylation of the CREB fragment with PKA. The GST CREB1-283 Δ (140-160) peptide could be phosphorylated with PKA, but did not bind CBP.

METHODS. The 78K and 58K CBP fragments were expressed using the bacterial expression vector pET-11d, modified to allow labelling by PKA¹⁶. For *a*, [^{32}P]CRE was preincubated at room temperature for 15 min with non-phosphorylated or phosphorylated CREB as described¹². Complexes were added to 25 ml buffer A and incubated with filters at room temperature for 2 h. Filters were washed for 60 min with three changes of buffer. For *b*, CREB was labelled with either PKA or CKII and incubated with the PKA-labelled 78K CBP fragment for 15 min at room temperature in buffer B containing 50 mM NaCl, 20 mM HEPES, pH 7.9, 5 mM MgCl_2 , 1 mg ml^{-1} BSA. Complexes were immunoprecipitated with anti-CBP antisera and protein A-Sepharose, washed, and analysed by SDS-PAGE and autoradiography. For *c*, recombinant CREB proteins were radiolabelled with PKA and [γ - ^{32}P]ATP. CREB1-160, CREB101-255 and CREB1-283 (Δ 140-160) were expressed and purified as GST fusion proteins. CREB1-341, CREB Δ Q1 and CREB Δ Q2 were purified by affinity chromatography on CRE columns. The CREB128-141 peptide was synthesized chemically. In *d*, GST-CREB proteins were bound to a glutathione column for affinity chromatography. The radiolabelled CBP fragment was incubated with each resin for 15 min at 4 °C in buffer B containing 50 mM NaCl, 20 mM HEPES, pH 7.9, 5 mM MgCl_2 , 1 mg ml^{-1} BSA. After washing the resin three times with buffer B, an equal volume of SDS loading buffer was added to resin, and eluted products were resolved by SDS-PAGE.



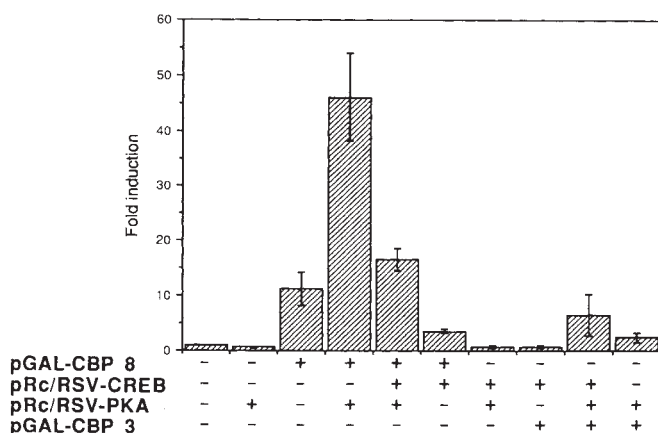


FIG. 4 GAL-CBP fusion protein activates expression of a GAL-CAT reporter *in vivo*. F9 teratocarcinoma cells were transiently transfected with the reporter plasmid pGAL-CAT and various combinations of the following plasmids as indicated: pGAL-CBP8, an expression plasmid encoding the DNA-binding domain of GAL4 fused to the N terminus of full-length CBP; pRc/RSV-CREB341, an expression plasmid encoding CREB341 (ref. 12); pRc/RSV-PKA, encoding the catalytic subunit of PKA (gift from R. Maurer); and pGAL-CBP3, encoding the DNA-binding domain of GAL4 fused to the N-terminal CREB-binding domain of CBP (amino acids 462–661). Transcriptional activity is reported as fold-induction compared to that with the reporter pGAL-CAT alone. Mean values and standard errors are shown; 4–5 independent experiments were done for each condition.

METHODS. Cells were transiently transfected by the calcium phosphate precipitation method¹⁷. pGAL-CBP8 was constructed by cloning cDNAs corresponding to amino acids 1–400 from CBP-2.8 and amino acids 401–2,463 from the mouse cDNA into the *Xba* restriction enzyme digestion site of pRSV-GAL (gift from M. Gilman). pGAL-CBP3 was constructed by subcloning the cDNA corresponding to amino acids 462–661 from CBP-0.8 into the *Xba* site of pRSV-GAL. The pGAL4-CAT reporter was the pGAL4/E1b TATA plasmid described in ref. 18. All transfections were normalized for the amount of plasmid and Rous sarcoma virus (RSV) promoter by the addition of the RSV plasmid (Invitrogen) and Bluescript plasmid (Stratagene). CAT (chloramphenicol acetyl transferase) was assayed as described previously and values were normalized to the amount of protein used in each assay.

induced genes. These processes may be directed by interaction with the zinc-finger or glutamine-rich domains that are shared with other transcriptional proteins. □

Received 4 May; accepted 31 August 1993.

- Montminy, M. R., Sevarino, K. A., Wagner, J. A., Mandel, G. & Goodman, R. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6682–6686 (1986).
- Short, J. M., Wynshaw-Boris, A., Short, H. P. & Hanson, R. W. *J. biol. Chem.* **261**, 9721–9726 (1986).
- Comb, M., Birnberg, N. C., Seasholtz, A., Herbert, E. & Goodman, H. M. *Nature* **323**, 353–356 (1986).
- Goodman, R. H. *A. Rev. Neurosci.* **13**, 111–127 (1990).
- Hoeffler, J. P., Meyer, T. E., Yun, Y., Jameson, J. L. & Habener, J. F. *Science* **242**, 1428–1432 (1988).
- Gonzalez, G. A. *et al. Nature* **337**, 749–752 (1989).
- Gonzalez, G. A. & Montminy, M. R. *Cell* **59**, 675–680 (1989).
- Berger, S. L. *et al. Cell* **70**, 251–265 (1992).
- Gonzalez, G. A., Menzel, P., Leonard, J., Fischer, W. H. & Montminy, M. R. *Molec. cell. Biol.* **11**, 1306–1312 (1991).
- Lee, C. Q., Yun, Y., Hoeffler, J. P. & Habener, J. F. *EMBO J.* **9**, 4455–4465 (1990).
- Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci.* **80**, 1194–1198 (1983).
- Walton, K. M., Chirvia, J. C., Rehfuess, R. P., Lochner, J. E. & Goodman, R. H. *Molec. Endocr.* **6**, 647–655 (1992).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Reichlin, M. *Meth. Enzym.* **70**, 159–165 (1980).
- Lamb, N. J. C., Fernandez, A., Watrin, A., Labbe, J. C. & Cavadore, J. C. *Cell* **60**, 151–165 (1990).
- Stofko-Hahn, R. E., Carr, D. W. & Scott, J. D. *FEBS Lett.* **302**, 274–278 (1992).
- Seed, B. & Sheen, J. Y. *Gene* **67**, 271–277 (1988).
- Lillie, J. W. & Green, M. R. *Nature* **338**, 39–44 (1989).

ACKNOWLEDGEMENTS. We thank R. Maurer, M. Gilman and M. R. Green for plasmids, R. Cone for the thyroid cDNA library, and J. Scott for help with the binding assays. Supported by grants from the NIH.

Molecular basis of crossreactivity and the limits of antibody–antigen complementarity

Jairo H. Arevalo*, Michael J. Taussig† & Ian A. Wilson*‡

* Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

† Laboratory of Structural Studies, Department of Immunology, The AFRC Babraham Institute, Babraham, Cambridge CB2 4AT, UK

Two major unanswered questions concerning the specificity of antibodies are: how do structurally different antigens bind with high affinity to the same antibody, and what are the limits of the antibody combining site complementarity and flexibility that contribute to such crossreactivity? We report here a comparative analysis of the X-ray structures of five conformationally different steroids in complex with the Fab' fragment of an anti-progesterone antibody DB3 at 2.7 Å. This antibody is unable to complement completely the shape of the hydrophobic antigen so that crossreactivity occurs with other ligands without major structural rearrangements of the binding site. Antigen specificity can be explained through conserved interactions of DB3 with the steroid D-ring, whereas some of the crossreactivity is realized through different binding orientations of the steroid skeleton that place the A-ring into alternative pockets on the antibody surface. The restricted gene usage of the VGAM3.8 family in the generation of anti-progesterone monoclonal antibodies^{1,2} may be explained by the specific interaction of VH hallmark residues with the steroid D-ring. This first detailed structure of steroid interactions with a protein could be applied to the understanding of general mechanisms of steroid recognition as well as in the design of specific binding sites for small hydrophobic ligands.

DB3 is a monoclonal antibody that binds progesterone with high affinity ($K_d \sim 1 \times 10^9 \text{ M}^{-1}$) and crossreacts with a small group of progesterone-like molecules with nanomolar affinities^{3–5}. DB3 can terminate pregnancy in mice, rats and ferrets, when administered shortly after mating⁴. The structural consequences of binding different high-affinity ligands to the same antibody combining site were analysed by X-ray crystallographic methods to understand the molecular basis of antibody–antigen crossreactivity and specificity^{6,7}. Structures of the Fab' fragment of an anti-progesterone antibody DB3 in complex with four progesterone-like steroid haptens (progesterone-11 α -hemisuccinate, 5 α -pregnene-3 β -hemisuccinate, aetiocholanolone and 5 β -androstane-3,17-dione) were determined by molecular replacement and difference Fourier methods using as an initial model the coordinates of the Fab' DB3–progesterone complex⁸ (Brookhaven accession code 1DBB; Table 1).

DB3 binds three conformationally distinct groups of progesterone-like steroids which differ at carbon C5 (unsaturated, 5 α and 5 β) in which substantial variation occurs in the position of the A-ring relative to the B, C and D rings of the steroid^{9,10}. A superposition of the five complexes reveals that the general shape of the DB3 combining site is highly conserved and that the steroid ligands adopt two distinct modes of binding. We term these modes 'syn' and 'anti' (Figs 1 and 2) and they are defined by the orientation of the steroid β -face (on which the two methyl groups in progesterone are located), relative to Trp^{H50}.

The syn binding orientation is characteristic of ligands with unsaturated C5 (sp²); progesterone, progesterone-11 α -hemisuccinate) or α substitutions at C5 (5 α -pregnene 3 β -hemisuccinate;

‡ To whom correspondence should be addressed.

Fig. 1 and Table 1). In these ligands, the coplanar (5α) or partially inclined ($\sim 45^\circ$ for sp² C5) A-rings fit into a cleft or hollow formed by His^{L27d} and Trp^{H100} on the α -face of the steroid and Val^{L94} on the β -face. The methyl positions, C18 and C19, face Trp^{H50}, and C11 (the linker position in the original immunogen) is towards the exterior of the combining site where it is accessible to solvent.

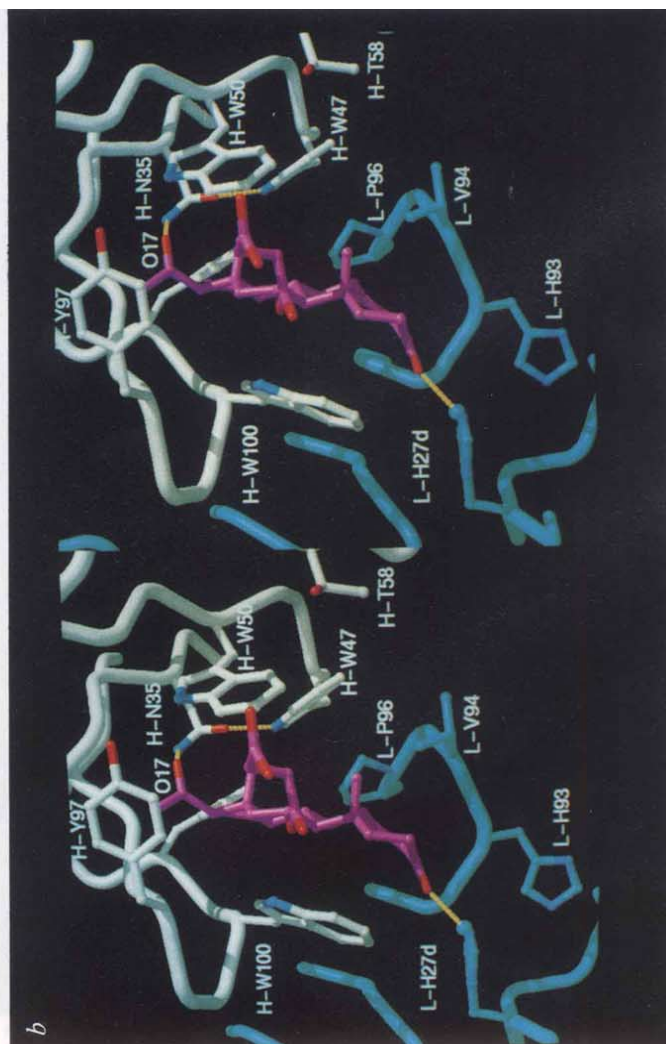
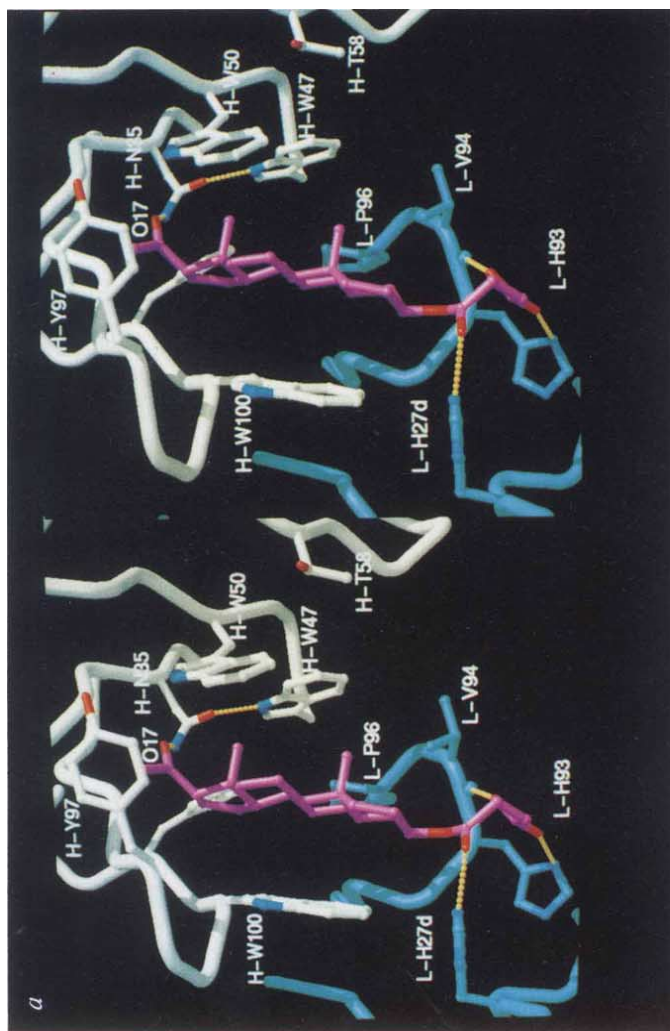
The *anti* orientation is characteristic of derivatives with 5β substitutions (aetiocolanolone and 5β -androstane-3,17-dione). In this family, the A-ring is perpendicular to the rest of the steroid skeleton (B, C and D rings^{9,10}). In DB3, the A-ring is accommodated in a different pocket, formed by Trp^{H50} and Val^{L94} on the α - and β -faces of the steroid, respectively (Fig. 2). As a result of an inversion of the steroid orientation, the methyl groups now make contacts with Trp^{H100}, whereas C11 is totally buried in the pocket.

In both conformations, the A-ring is held in hydrophobic pockets, where the different C3 functional groups (keto, hydroxyl and hemisuccinate) are stabilized by hydrogen-bond interactions. More precisely, in the *syn* conformation, the keto or hemisuccinate linker groups at C3 are involved in hydrogen bonds with CDR-L1 residue His^{L27d} (N ϵ 2) or with the CDR-L3 loop residues His^{L93} (N δ 1) and Val^{L94} (N) (Fig. 1). In the *anti* conformation, the interactions of the keto and hydroxyl groups are mediated in a different pocket through hydrogen bonds with a bound water molecule (Wat1; Figs 1 and 2).

The common interactions that are relevant to steroid recognition centre on the D-ring and the DB3 heavy chain. In the five steroid complexes, the D-ring is deeply embedded in a hydropho-

bic cavity formed by Trp^{H50}, Tyr^{H97} and Trp^{H100}, which are stacked in a perpendicular fashion¹¹, and Phe^{H100b} at the bottom of the pocket. These specific interactions with the D-ring are maintained despite the two different steroid docking orientations and a shorter substituent (a keto group) at C17 in the *anti* conformation steroids compared to the methyl-keto group of the *syn* progesterone-like steroids (Figs 1, 2 and Table 1). Hydrogen bonds are made with Asn^{H35} NH δ and the functional keto groups at positions C17 or C20 (Figs 1 and 2). The B and C steroid ring faces are sandwiched between the indoles of Trp^{H50} and Trp^{H100}, which delimit the groove where the steroid skeleton docks (Figs 1 and 2). Most of the A- and B-ring interactions are at the interface of the light and heavy chain where Val^{L94} divides the combining site into two pockets for the alternative α and β configurations at C5 (Figs 1 and 2). Thus, the antibody provides specific hydrogen bonds for the only two functional groups of the steroid and forms a complementary molecular surface with the non-polar skeleton.

The DB3 antibody has two different conformations defining the free ('closed'; IDBA) and progesterone-bound ('open'; IDBB) states⁸. No large changes are seen among the various steroid-bound forms; the backbone (root-mean-square difference, r.m.s.d. ~ 0.35 Å) and side chains (r.m.s.d. ~ 0.75 Å) throughout the pocket have a small degree of flexibility that defines the limits of induced fit complementarity for this antibody^{12,14}. The regions with the highest degree of flexibility are the indole of Trp^{H100}, which can move by up to 25° , and residues L93–L96 in the CDR-L3 where the backbone (r.m.s.d. ~ 0.6 Å) and side-chain atoms (r.m.s.d. ~ 0.9 Å) contri-

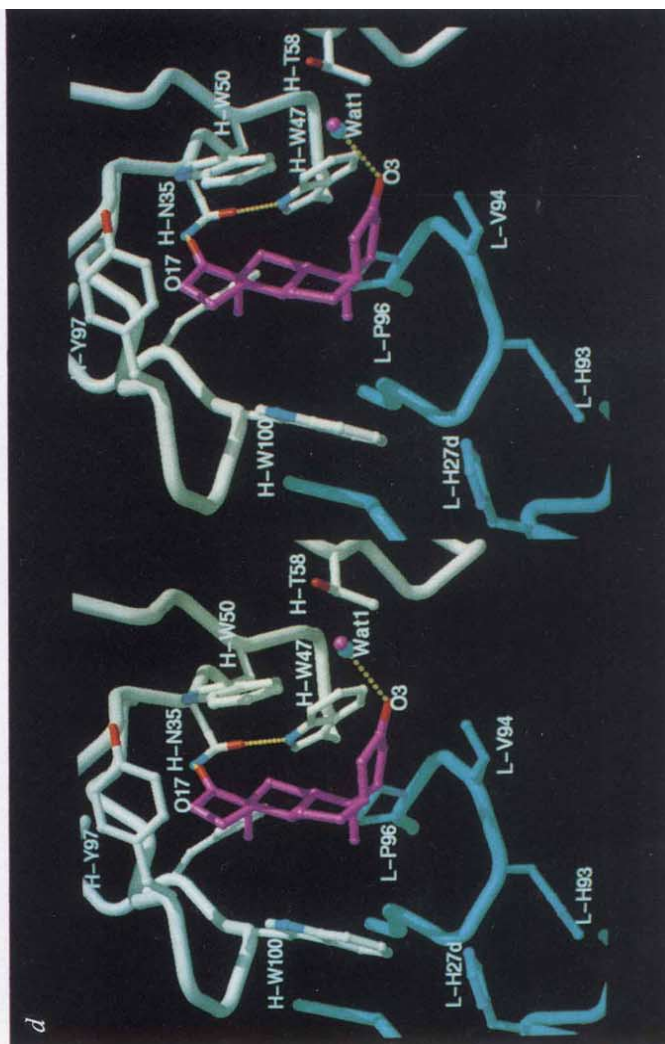
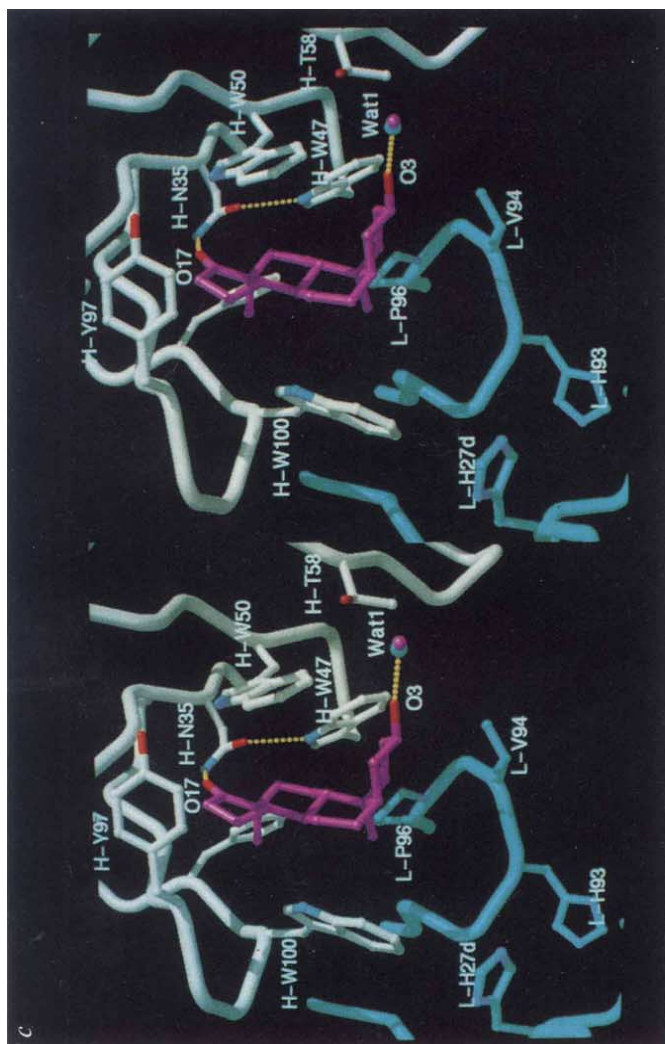


bute to the fit for different ligands (Figs 1 and 2).

A measure of the tightness of fit for each ligand in the combining site reveals where the steroid skeleton has the highest degree of complementarity with the antibody. Measurements of the maximum radius from each ligand atom intercepting the extended surface^{15,16} of the combining site show that the tightest fit corresponds to the keto groups at positions C17 or C20 as well to other atoms in the D-ring; these atoms also display the highest number of van der Waals interactions. Based on these structural findings and supported by binding assays^{3,4,8} we propose that the specificity of this antibody-antigen recognition is centred on the D-ring and the hydrogen bonds to O17 or O20.

The comparative looseness of fit at the core of the steroid skeleton contrasts with the specific A- and D-ring interactions and may be associated with the linker being at the 11 α position in the original immunogen³. The 11 α -hemisuccinate moiety is accessible to solvent and indicates where the carrier protein could have been positioned for hapten presentation. The indole sandwich formed by Trp^{H50} and Trp^{H100} can accommodate alternative orientations of the ligand core and may limit the complementarity that can be achieved for steroids with the DB3 antibody. Somatic mutation from the original germline V_H and V_L genes did not generate antibodies with substantially different residues in the binding site and their affinities for progesterone were not considerably improved^{1,2,4}. It is not clear to what extent variable gene restriction and diversity is determined by the nature of the hapten or by the chemistry of its conjugation to

FIG. 1 Steroid binding site showing crossreactive ligands. Stereo diagrams (in identical orientations) of the DB3 backbone and side chain atoms in contact with the four steroid haptens (magenta). The DB3 light (cyan) and heavy (white) chain are colour coded to highlight the distribution of interactions between domains. 5 α -pregnane-3 β -ol-hemisuccinate (a) and progesterone-11 α -ol-hemisuccinate (b) are examples of the syn binding orientation, where aetiocholanolone (c) and 5 β -androsterane-3,17 dione (d) exemplify the antibinding orientation. The functional keto group at the D ring C17 or C20 is stabilized by a conserved hydrogen-bond interaction with Asn^{H35} (NH δ). The Asn^{H35} (Q δ) hydrogen bonds with the N ϵ of Trp^{H47} orienting the Asn side chain amide such that a hydrogen-bond acceptor is favoured at C17 or C20. The C3 functional groups of the syn docking conformation hydrogen bond with His^{L27d} (CDR-L1) and residues from the CDR-L3, His^{H93} and Val^{L96}, whereas in the anti conformation the hydrogen-bond interactions are made with a water molecule Wat1 (blue), which is close to Thr^{H58} (3.7 Å). Specific recognition of the A- and D-rings through hydrogen bonds is clearly seen for all the complexes. This picture was produced on a Crimson Silicon Graphics using the program MIDAS-PLUS³¹.



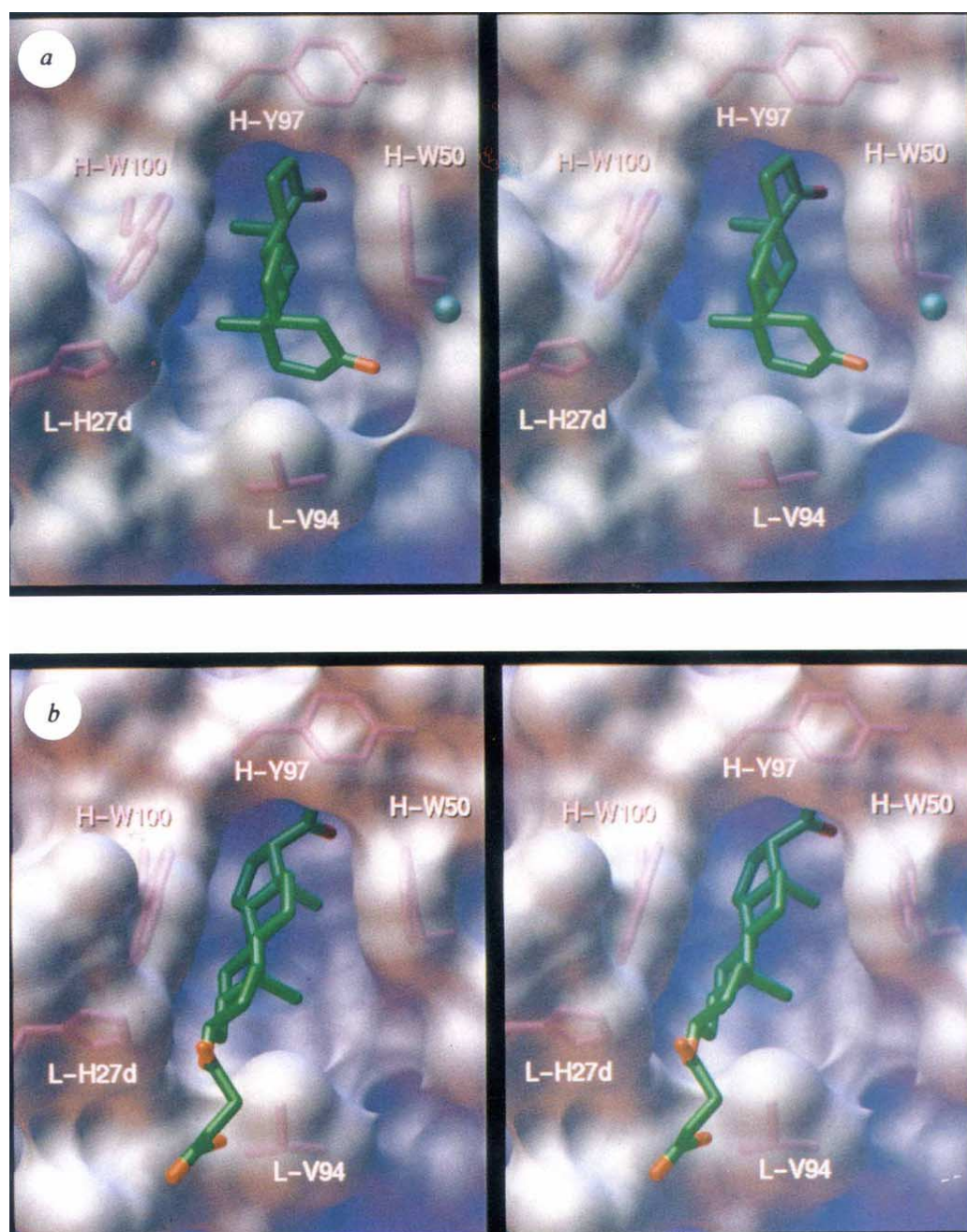


FIG. 2 DB3 binding site surface showing crossreactive ligands. A surface representation of the binding sites for the *anti* (a) 5 β -androstane 3,17-dione and *syn* (b) 5 α -pregnane-3 β -ol-hemisuccinate conformations. The Fab' DB3 side chains in contact with the steroid are in magenta. The interaction of the different ligands A-rings with the two alternative pockets in the antibody is illustrated and shows how the two different orientations of the steroid are aligned to conserve the specificity of the D-ring interactions in the interior of the pocket. The surface was calculated with the program PQMS³² using a 1.5 Å probe radius and van der Waals radii defined in ref. 33. The molecular tube representation was produced with TUBES written by Y. Chen and A. Olson and rendered in the AVS environment³⁴.

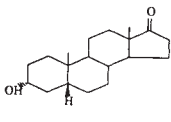
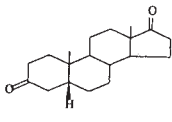
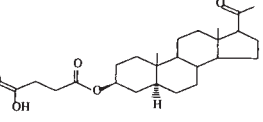
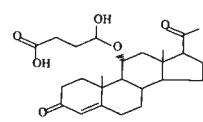
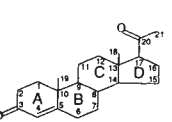
the protein carrier^{17,18}.

The murine anti-progesterone response identified germline antibodies with a remarkably high affinity for progesterone ($K_a \sim 10^9 \text{ M}^{-1}$) that are severely restricted to a small gene family for the light ($V_{\kappa} 5.1$) and heavy (VGAM3.8) chains². Similar antigen-gene associations have been found for other antibodies^{19,20}, including the NQ10/12.5 2-phenyloxazolone Fab complex²¹ and the phenylarsonate-Fab complex^{22,23}. The significance of the restricted germline V-region genes selected for the heavy chain of anti-progesterone antibodies^{1,2} is revealed by the important interaction of the VGAM3.8 hallmark residues Asn^{H35} and Trp^{H50} with the steroid. These residues are directly involved in the molecular constraints associated with hapten recognition and specificity. More precisely, Asn^{H35} defines the

nature of the hydrogen-bond acceptor at C17 or C20, whereas Trp^{H50} and Trp^{H100} form the predominant interactions with the B- and D-rings (Figs 1 and 2).

Thus, DB3 specificity centres on a common set of interactions defining the minimal pharmacophoric elements for binding and suggests a new path in the *de novo* design of novel molecules that mimic progesterone. The paucity of functional groups in progesterone makes it difficult to engineer a binding pocket with high specificity that avoids some crossreactivity with related ligands. That proteins in general have difficulty in the fine discrimination of steroids is reflected in the known crossreactivities observed in steroid receptors^{24,25} and in the binding orientation variability of steroid-metabolizing enzymes^{26,27}. In spite of the challenges posed to the immune system in recognizing a

TABLE 1 Anti-progesterone crossreactive ligands

					
	Aetiocholanolone	5β-Androstane 3,17-dione	5α-Pregnane-20-one 3β-hemisuccinate	Progesterone 11α-hemisuccinate	Progesterone
Affinities IC ₅₀	21 nM	8 nM	2 nM	0.36 nM	1 nM
Binding conformation	anti	anti	syn	syn	syn
Fractional ligand buried	91%	88%	82%	81%	89%
Ligand buried area (Å ²)	227	223	288	291	241
Antibody buried area (Å ²)	304	294	348	353	270
Number of van der Waals and hydrogen-bond interactions	27 (2)	20 (2)	60 (4)	68 (2)	59 (2)
Refinement resolution	6–2.7 Å	6–3.0 Å	6–2.9 Å	6–2.7 Å	6–2.7 Å
*R-value (%)	21.4%	21.1%	20.5%	21.8%	21.4%
r.m.s. Δ bond length	0.016 Å	0.018 Å	0.019 Å	0.017 Å	0.018 Å
r.m.s. Δ bond angles	3.90°	3.95°	3.81°	3.83°	3.90°

The different steroids studied all bind DB3 in the nanomolar range. The IC₅₀ relative affinities were determined by enzyme-linked immunosorbent assay (ELISA) in which binding of progesterone alkaline phosphatase was inhibited by free steroid. The hydrogen bonds (in parenthesis) and van der Waals interactions are reported for all the structures determined. The reduced number of interactions for the 5β derivatives is consistent with their slightly lower affinity. Fab' crystals suitable for structure determination were grown by vapour diffusion experiments in 1.7 M ammonium sulphate, 0.15 M imidazole citrate buffer, pH 7.4, using as additives 0.002% (w/v) PEG 20,000, 1% (v/v) ethanol and with Fab at 5 to 10 mg, 0.1 M ml⁻¹ (ref. 5). All the cocrystal complexes are isomorphous with the unliganded and progesterone-bound Fab' complex, and belong to the hexagonal space group *P* 6₃ 2 2 with unit cell dimensions *a* = *b* = 134.76 Å, *c* = 124.21 Å (ref. 5). The structures were solved by molecular replacement and difference Fourier methods using as starting model the progesterone–Fab' complex coordinates (1DBB; ref. 8). Refinement was done using XPLOR2.1 (ref. 28) and manual rebuilding was aided with FRODO²⁹. Current *R*-values for the five structures range from (20.5–21.8%) at a resolution of 2.7–3.0 Å without including a solvent model and with average *B*-values for all atoms ranging from 19.7 Å² (5β-androstane 3,17-dione) to 29.3 Å² (aetiocholanolone). A complete description of the refinement and structure determination will be presented in detail elsewhere. Fab' residue numbering follows ref. 30 throughout.

$$* R\text{-value} = \frac{\sum_{hkl} ||F_o| - |F_c||}{\sum_{hkl} |F_o|}$$

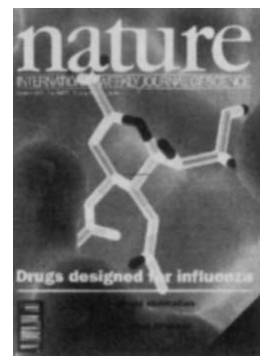
largely apolar and functionally inert steroid skeleton, anti-progesterone antibodies have found a way to create a high-affinity binding site, but which is limited in its complementarity to the antigen. □

32. Connolly, M. L. *Science* **221**, 709–713 (1983).
33. Gelin, B. R. & Karplus, M. *Biochemistry* **18**, 1256–1268 (1979).
34. Upson, C. et al. *IEEE Comput. Graph. Applic.* **9**, 30–42 (1989).

ACKNOWLEDGEMENTS. We thank A. Feinstein for initial efforts in the development of the anti-progesterone antibody programme, E. A. Stura for defining the DB3 crystallization conditions, R. Lerner, B. Heap, T. Yeates, D. Filman, J. Johnson, D. Jewell and R. Stanfield for constant encouragement and support throughout the project, C. A. Hassig for computational assistance, S. Humphreys for the binding analyses and M. Pique for preparation of the figures. This work has been supported by the NIH (I.A.W.), TSRI graduate program (J.H.A.), US Department of Energy and SSRL, Ortho Pharmaceutical (M.J.T.) and Johnson and Johnson (M.J.T.). Coordinates have been deposited in the Brookhaven Data Bank and are available from Prelease.

ERRATUM

THIS cover picture of the 3 June issue (*Nature* **363**, issue 6428; 1993) depicted an inhibitor of influenza virus bound to its target, as described on page 418 in an article by M. von Itzstein et al. In the caption to the cover picture given on the Contents page, an acknowledgement was omitted to the authors of the program used to generate this graphic. The program was GRASP and its authors were A. Nicholls and B. Honig of the Department of Biochemistry, Columbia University, New York. □



Received 7 June; accepted 23 August 1993.

1. Deverson, E., Berek, C., Taussig, M. J. & Feinstein, A. *Eur. J. Immun.* **17**, 9–13 (1987).
2. Sims, M. J., Krawinkel, U. & Taussig, M. J. *J. Immun.* **149**, 1642–1648 (1992).
3. Wright, L. J. et al. *Nature* **295**, 415–417 (1982).
4. Ellis, S. T. et al. *J. Endocrin.* **118**, 69–80 (1988).
5. Stura, E. A. et al. *Immunology* **62**, 511–521 (1987).
6. Landsteiner, K. *The Specificity of Serological Reactions* (Dover, New York, 1962).
7. van Regenmortel, M. H. V. *Immun. Today* **10**, 266–272 (1989).
8. Arevalo, J. H., Stura, E. A., Taussig, M. J. & Wilson, I. A. *J. molec. Biol.* **231**, 103–118 (1993).
9. Duax, W. L. & Norton, D. A. *Atlas of Steroid Structures* Vol. 1 (Plenum, New York, 1975).
10. Griffin, J. F., Duax, W. L. & Weeks, C. M. *Atlas of Steroid Structures* Vol. 2 (Plenum, New York, 1984).
11. Burley, S. K. & Petsko, G. A. *Science* **229**, 23–28 (1983).
12. Wilson, I. A. & Stanfield, R. L. *Curr. Opin. struct. Biol.* **3**, 113–118 (1993).
13. Stanfield, R. L., Takimoto-Kamimura, M., Rini, J. M., Profy, A. T. & Wilson I. A. *Structure* (in the press).
14. Rini, J. M., Schulze-Gahmen, U. & Wilson, I. A. *Science* **255**, 959–965 (1992).
15. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 376–400 (1971).
16. Bohacek, R. S. & McMartin, C. J. *med. Chem.* **35**, 1671–1684 (1992).
17. Mudgett-Hunter, M., Anderson, W., Haber, E. & Margolies, M. N. *Molec. Immun.* **22**, 477–488 (1985).
18. Taussig, M. J. et al. *Immunology* **72**, 471–480 (1991).
19. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. *Sequences of Proteins of Immunological Interest* 5th edn (National Institutes of Health, Bethesda, 1991).
20. Berek, C. & Milstein, C. *Immunol. Rev.* **96**, 23–41 (1987).
21. Alzari, P. M. et al. *EMBO J.* **9**, 3807–3814 (1990).
22. Parhami-Seren, B., Kussie, P. H., Strong, R. K. & Margolies, M. N. *J. Immun.* **150**, 1829–1837 (1993).
23. Lascombe, M.-B., Alzari, P. M., Poljak, R. J. & Nisonoff, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9429–9433 (1992).
24. Evans, R. M. *Science* **240**, 889–895 (1988).
25. Baulieu, E.-E. *Science* **245**, 1351–1357 (1989).
26. Strickler, R. C., Covey, D. F. & Tobias, B. *Biochemistry* **19**, 4950–4954 (1980).
27. Luu-The, V. et al. *Biochemistry* **30**, 8861–8865 (1991).
28. Brunger, A. T. *X-PLOR* version 2.1 (Yale University, 1990).
29. Jones, A. T. *J. appl. Crystallogr.* **11**, 268–272 (1978).
30. Kabat, E. A. & Wu, T. T. *J. Immun.* **147**, 1709–1719 (1991).
31. Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. J. *molec. Graphics* **6**, 13–27 (1988).

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.

New product watch

New on the market this week — a DNA sequencer that delivers typical read lengths of 800 bases per sample, software for HPLC users and a solvent recycling system designed to reduce costs.

A NEW low-cost version of Pharmacia's BIAcore biosensor-based analytical instrument for **real-time, label-free interaction analysis**, called the BIAlite, is now available (*Reader Service No. 100*). BIAlite is a single-channel, manual version of BIAcore, which has four-channels and is automated.



BIAlite — a low-cost version of Pharmacia's BIAcore biosensor.

This small, computer-controlled instrument is based on the same precision fluidics and surface plasmon detection as BIAcore, yielding rapid interaction analysis results such as relative binding patterns, active concentration in complex mixtures, kinetic rate constants including both on and off rate constants, and affinity constants by either Scatchard analysis or rate constants. Pharmacia says that typical interactions take between 3 and 10 min. User-definable sensor chips can be regenerated and are said to yield about 100 injections without variation.

The new **portable photosynthesis and transpiration measurement system** from ADC, the LCA4, features automatic water vapour and carbon dioxide control facilities (*Reader Service No. 101*). The LCA4 offers



ADC's portable photosynthesis and transpiration measurement system.

plant scientists a choice of parameters affecting rates of photosynthesis, metabolism and stomatal conductance. Other design features include the option of battery or mains operation, a selection of leaf chambers and a transducer kit that allows the user to design and build a chamber. Typical applications include investigating the effects of environmental changes such as air pollution and elevated CO₂ levels, and use in crop management and yield improvement schemes. Based on non-dispersive, infrared analysers, the system provides measurements of CO₂ in the 0–1,600 p.p.m. or 0–3,000 p.p.m. range, and water vapour in the 0–70 millibar range. Data recording can be operated through the keypad, by a button on the leaf chamber, remotely, or automatically at pre-set intervals. The system offers unlimited data storage on replaceable memory cards, which the user can pre-programme, and outputs for interfacing with laboratory equipment, printers or computers.

The new 160-page 1993–1994 **catalogue** from Peptides International contains information and references for the company's hundreds of biologically active peptides, enzyme inhibitors, peptide-related building blocks and equipment (*Reader Service No. 102*). One of the catalogue's most useful features is its three indexes, including a grouping based on amino acid identity. The company specializes in peptides of exceptionally high purity, and especially structurally complex compounds such as charybdotoxin, iberiotoxin and omega agatoxin IVA.

The new Spectracube 1000 **imaging spectrometer** from Metax provides cell biologists and biochemists with an instrument that analyses the microscope image pixel by pixel (*Reader Service No. 103*). It simultaneously measures the visible-to-near-infrared spectrum of every pixel making up the image, providing a tool for 0.4–1 spectroscopic measurements at every point on the image plane. This makes it possible to analyse the content of an image by frequency, thereby defining and making clear the spectral characteristics of, for example, an image of a cell. Using the Spectracube for transmission spectroscopy, the microscope image — of a cell, for example — is displayed on a screen. The spectrum of each point can then be displayed by moving the cursor, or the data may be processed to display colour-coded spectral signatures of chemical components superimposed on the image of the cell. Data may also be stored for future processing, com-

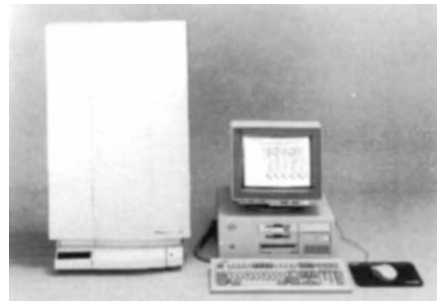


Metax's new imaging spectrometer analyses the microscope image pixel by pixel.

parison with other cells, or comparison with the same cell at a later date. The Spectracube may also be used for reflection spectroscopy, in, for example, the manufacture of integrated circuits and quality assurance procedures. The key to the instrument's versatility is said to be its accompanying software, which allows spectral images to be displayed, processed and stored on a personal computer. Each spectral image is a complete spectral cube of information, consisting of hundreds of stored images of a single scene, each at a distinct wavelength. From these data, the spectral cube can be displayed as a pseudo colour image, including a spectral display of any pixel.

■ Molecular biology miscellany

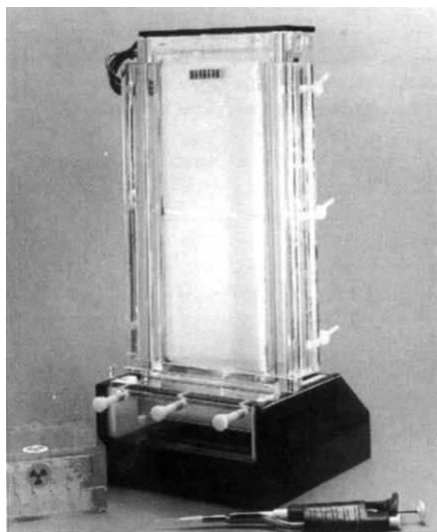
Sequence read lengths up to 800 bases or more are typical with Li-Cor's Model 4000L **automated DNA sequencer** (*Reader Service No. 104*). The Model 4000L utilizes infrared fluorescence sequencing technology and a 66-cm gel to yield single-base



Li-Cor's Model 4000L automated DNA sequencer — said to provide read lengths of 800–1,000 bases per sample.

resolution at base 800. For shorter DNA fragments, other gel sizes have run times as short as 1 h, Li-Cor says. Samples can be prepared using Sanger dideoxy protocols with labelled primers or infrared dye-labelled dATPs. Cycle sequencing protocols are also available. DNA fragments labelled with an infrared fluorophore are excited with a solid-state laser diode and the infrared emission detected with a solid-state, silicon avalanche photodiode. Li-Cor says that solid-state infrared technology increases reliability and throughput while reducing the cost of the instrument and operating costs. Electrophoresis results are presented in an autoradiogram-like image format. The Base ImagIR software provides automatic base calling with a stated accuracy for single-stranded DNA of 99 per cent to 800 bases.

The new Polar Bear **isothermal electrophoresis system** from Owl Scientific allows researchers to run sequencing, single-stranded conformational polymorphism, mobility shift assays and other applications



Eliminate the need for specialized gel matrices, toxic chemicals, fans, dry ice or a coldroom with the Polar Bear system.

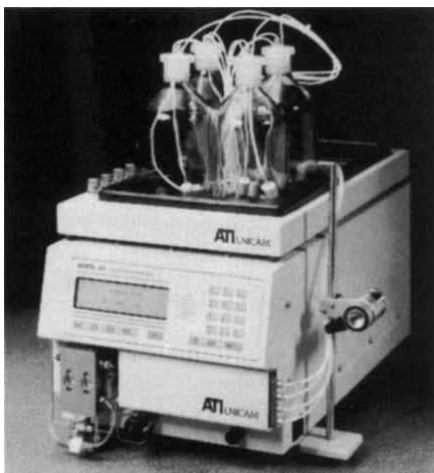
with precise temperature control (*Reader Service No. 105*). This 20 × 45-cm system can be used with constant temperature water or coolant systems over a wide temperature range for high-resolution detection of genetic mutations, polymorphisms, DNA-binding proteins, dinucleotide repeats and point mutations of PCR products. Control is provided by the system's temperature exchange system, which incorporates an alumina ceramic interface for optimal heat transfer. Maximum contact between the temperature control chamber and the gel plate ensures uniformity of temperature across the gel. The system features a safety interlocking lid and a "leak-free" clamping system designed to provide even pressures across the entire gel surface. It is supplied with gel plates, spacers, a 39-tooth sharktooth comb, tubing and stopcocks.

Thermal prints can be ready in less than 5 s with the new **Visionary gel documentation station** from Fotodyne (*Reader Service No. 106*). The benchtop darkroom, which has been designed to cut the cost of gel photography, allows the user to take photographs of gels, blots, TLC plates or other data without leaving the bench. Images can be previewed on the video monitor before printing. The image integration facility ensures the capture of low-light images, such as gels stained with ethidium bromide.

DNASTAR's Lasergene biocomputing software for Microsoft Windows will be available from January 1994 (*Reader Service No. 107*). This new **molecular biology software** package for Windows is compatible with the company's Macintosh version of Lasergene, and so documents and files created on one computer can be used on the other, without modification. The capabilities of Lasergene include sequence assembly, database searching, multiple and pairwise sequence alignment, protein analysis, graphical sequence mapping and primer selection and design. The Windows version of Lasergene, like the Macintosh version, is said to perform well on computer networks. Advanced versions of the software for integrated networks supporting both platforms simultaneously are under development. In addition to Lasergene, DNASTAR is offering Primer Select for both Macintosh and Windows. The program provides computer-aided oligonucleotide design for the researcher using polymerase chain reaction or primer-based sequencing techniques. It recommends optimal primer pairs and experimental strategies that are most likely to work in the test tube, avoiding problems such as unequal melting temperatures and secondary structure interference.

■ HPLC

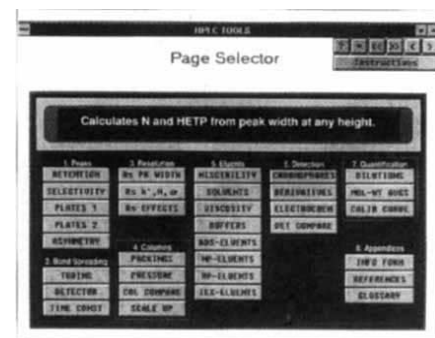
The Unicam Crystal 200 pump is a **modular solvent delivery system for HPLC** that offers single solvent, binary and quaternary options (*Reader Service No. 108*). It can store up to ten methods in nonvolatile



Unicam's Crystal 200 HPLC pump has a modular make-up.

memory; these can be sequenced together when performing automatic unattended analysis. Online solvent gassing is offered by either helium or vacuum. In addition, the single solvent pump can be upgraded to the full quaternary isocratic/ternary gradient system. Designed to span narrow bore to analytical flow rates, the pump has seven design features that are intended to ensure an accurate and reproducible flow and mix of solvents. The pump head features a piston wash and sapphire piston chambers, which enable the operator to observe operation of both the piston and check valve. Although the pump has a standalone capability, it can also be controlled from a computer running Windows-based control software.

HPLC Calculations Assistant and Reference Tools is the new Windows-based **HPLC software program** from Savant Audiovisuals that should provide a useful tool to chromatographers needing to perform various HPLC calculations (*Reader*



Savant's software — designed to make light work of HPLC calculations.

Service No. 109). Developed by Dr Dennis Saunders of the Unocal Science and Technology Division, the program features a master Page Selector, which allows users to choose any of 33 built-in functions. It contains individual calculators related to peaks, band spreading, resolution, columns, eluents, detection and quantification. For example, calculators in the "Peaks" section determine retention time, peak velocity, retention volume, selectivity, plates, HETP and asymmetry. When in operation, users may input parameters that are specific to their own experiments. The software also includes two onscreen plotting devices, eight reference charts and tables, printout resources, a bibliography and a glossary of HPLC terms.

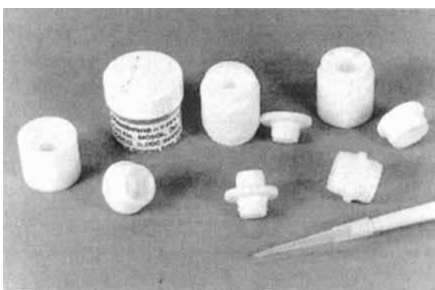
For all chromatographers seeking an HPLC pump that satisfies the demand for biocompatibility, inertness and an absence of metal parts that can come into contact with solvents, Merck may have the answer with its new **L6220 inert HPLC pump** for biochromatographic separations (*Reader Service No. 110*). Merck says that inert polymers (such as pressure-resistant poly ether ether ketone) and ceramics have been used to ensure that the pump is free of metal. The

pump's system control functions are useful in biochromatographic applications where gradients are used or where valves, accessory pumps or other instruments must be controlled automatically. For central control, the D6000 HPLC manager software can be used, while a gradient accessory can be added to form a ternary low-pressure gradient pump eliminating the need for a second pump during gradient elution. The pressure range extends up to 275 bars, which makes the pump adaptable for applications in analytical HPLC, biochromatography and ion chromatography. A special piston rinsing mechanism that can be used with flowing and stationary washing liquids eliminates the danger of wear from precipitated buffers.

■ Laboratory gadgetry

BioNeb is the new **disruption system** from Hoefer Scientific that was invented and developed at Indiana University by Dr Stefan J. Surzcki, Dr Robert K. Togasaki and Mr Masahiko Kitayama (*Reader Service No. 111*). The device offers researchers a way to disrupt cells and polymer molecules such as DNA for analysis. It operates using a low-pressure flow of gas to generate shear force during liquid nebulization. Hoefer says that this method of disruption creates no heat and is gentle enough to open cells for organelle preparation, or it can be adjusted to break even the toughest cells.

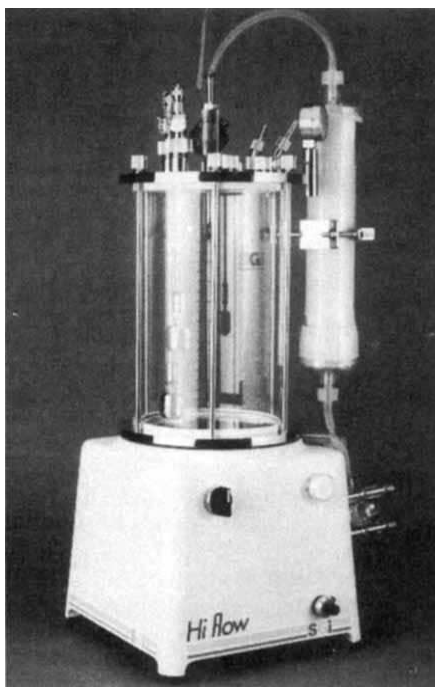
Sialomed introduces a new **microdialysis system** that comes with its own built-in magnet and which is designed for the rapid purification of biological samples in microlitre quantities (*Reader Service*



Sialomed's microdialyser comes with its own built-in magnet.

No. 112). The built-in magnet further provides a better stirring of the buffer as well as spinning of the microdialyser. The system is made of Teflon and comes in different sizes: 10, 20, 50, 100, 200, 500, 1,000 and 1,500 μ l. Precut ultrafiltration membrane sizes include: 0.1 K, 0.5 K, 1 K, 2 K, 5 K, 10 K, 25 K, 50 K, 100 K and 300 K. It can be used for the preparation of biological samples including electro dialysis, electroconcentration, electroelution, and the removal of primers after the polymerase chain reaction.

The SGI Hi-Flow laboratory **crossflow filtration unit** from LH Fermentation is designed for use where bench-scale filtration using hollow-fibre cartridges is required



For continuous/batch mode processing — see the Hi-Flow crossflow filtration unit.

(*Reader Service No. 113*). The unit can be autoclaved and is semi-automatic, allowing it to be used in continuous as well as batch mode. It is supplied with a 4- or 7-litre vessel, a built-in peristaltic pump for circulation and level control enabling operation of an external pump where large volumes are to be processed. The unit can be used with many of the currently available cartridges or can be supplied as a complete system. Applications include concentration, clarification, diafiltration and sterilization of media.

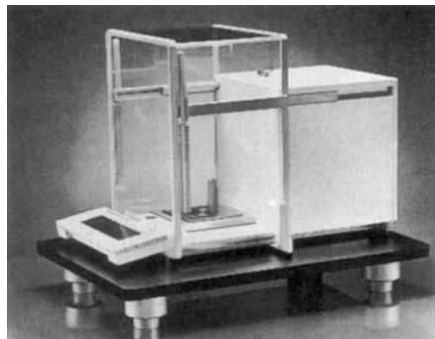
To reduce solvent costs and the danger of toxic waste disposal, Life Sciences International is distributing the Varisol 12 litre **solvent recycler**, which is manufactured by



Waste not want not with solvent recycling.

Shandon (*Reader Service No. 114*). This compact, floorstanding model can be used to recycle commonly used stable solvents including xylene and alcohol, and should be of interest to those working in histopathology laboratories. LSI says that the device takes only 4 h to recover solvents with a boiling point between 40 °C and 200 °C and has a stated recovery rate of up to 96 per cent, depending on the contamination grade. Once set, the Varisol can be left to run automatically. An integral stand with drip tray and a simple tipping mechanism assists in cleaning, filling and servicing. For safety, the Varisol operates under normal atmospheric pressure and incorporates several over-temperature cut-out systems, on the outlet pipe for cooled solvent delivery and on the oil temperature to prevent overheating.

Precision Dynamics has introduced the new Mach I-AB clarifier **antivibration stand** that is designed to protect analytical balances from the effects of vibration (*Reader Service No. 115*). Measuring only 3 inches high and weighing only 40 lbs, the Mach I-AB is said to eliminate "scrambled" digital readouts caused by excessive vibration levels. The system consists of a compact, rigid worksurface that is supported



Steady as a rock with the Mach I-AB antivibration stand for precision balances.

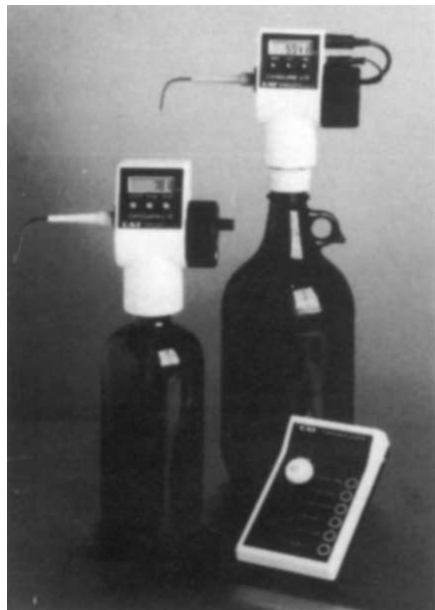
by four high-performance isolators — one located at each corner of the unit. These isolators are at the heart of the system; each includes a helical coil spring sealed inside a protective cast aluminium housing. When in use the isolators exhibit a 4.0 Hz natural frequency — said to protect analytical balances against most vibration conditions.

With the new Haake **digital-controlled immersion heater**, the pre-selected temperature can be maintained to within a few hundredths of a degree (*Reader Service No. 116*). In addition, the immersion circulators maintain the exact temperature digitally, circulate the liquid in the bath in order to obtain an even temperature distribution, and make unsupervised continuous operation possible by way of numerous safety elements.

■ Motorized means

Polyscience has introduced a new **motorized continuous digital burette** with a microprocessor-controlled controller for pre-

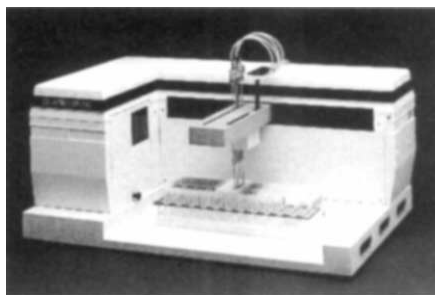
cise repetitive dispensing (*Reader Service No. 117*). The controller has a two-line alphanumeric display with such functions as dispensing pre-set volumes, repeat dosing and automatic set volume. The new motorized burette's auxiliary outlet can also be used with PC or TTL connection. This allows easy connection to a level sensor, pH meter or a microcontroller or PC. Both the manual and motorized burettes mount directly to the reagent bottle, have an operating range of 0.01 to 9999.99 ml with a stated precision of ≤ 0.1 per cent of the rated volume, and do not have valve mechanisms or syringe pistons to jam, clog or leak.



Polyscience's motorized liquid dispensing system.

Polyscience also says that because the maximum stored volume is 0.2 ml, the dead volume is kept to a minimum. The solid-state electronic memory is cumulative to 9999.99 ml and the six-digit LCD readout eliminates meniscus reading errors. The ceramic and Teflon construction is resistant to acid and bases.

Quatro Biosystems is looking to clinical laboratories as the primary market for its new SP-240 four-probe **robotic sample processor** (*Reader Service No. 118*). An enhanced version of the 200 series, the SP-240 offers a greater choice of probe functions, including a post-wash facility for sample and reagent transfers, and large reagent boats for higher volume processing. More flexibility in the choice of rack layouts has also been built in to the new design. In addition, an internal coating of Teflon on the probe is said to reduce carryover and improve reagent compatibility. For added security and reliability of results, an error recording and review option alerts the user to any faults that have occurred during processing. Other features include a four-hole wash trough and enhanced fluidic washing facilities



SP-240 four-probe robotic sample processor.

ties to improve both internal and external probe washing. The outer casing mechanics and electronics of the SP-240 have been upgraded to include new linear slide-mounted x and y axes. The instrument can be programmed using either a hand-held controller or a PC.

■ Separation modes

Avanti 30 is the new **refrigerated tabletop centrifuge** from Beckman designed for the simple and fast separation of small-volume samples (*Reader Service No. 119*). The Avanti has a small footprint, offers a choice of customized acceleration and deceleration profiles and performs at 30,000 r.p.m. with up to 64,000 g using a fixed-angle rotor. Other



Small-volume sample separation with Beckman's Avanti 30.

design features include microlitre sample volume, a tube capacity of up to 85 ml, an intuitive user interface, a brushless induction drive and a speed that is selectable in r.p.m. or r.c.f. Protection against rotor imbalance, over-temperature, over-speed and exposure to aerosols, a barrier ring and door interlock are standard safety features.

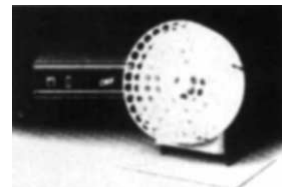
The Dynal MPC-9600 **magnetic particle concentrator** is designed for the rapid isolation of Dynabeads' products in up to 96 reaction tubes (0.2 ml) or up to 48 reaction tubes (0.5 ml) (*Reader Service No. 120*). The concentrator is constructed of disinfectant-proof material and contains stable rare-earth elements (neodymium-iron-boron permanent magnets) that are intended to separate

the magnetic beads. Nucleic acids bound to the Dynabeads are separated from the reaction volume and attracted to the sides of the microtubes. The MPC-9600 is suitable for use in RT-PCR and solid-phase DNA sequencing applications when using Dynabeads Oligo (dT)₂₅ or Dynabeads M-280 streptavidin products. For solid-phase DNA sequencing, the MPC-9600 allows the PCR, template preparation and DNA sequencing steps to be performed in the same microtube. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

CELLMIXER LABORTECHNIK FRÖBEL



sturdy, reliable, versatile

speeds: 10...100 rpm CMV-1-100
1...15 rpm CMV-1-15

angle: from horizontal to vertical

*optional: up to 70°C/158°F, 100% rel. humid.
Type CMV-2-100 or CMV-2-15 for use in incubators, with control unit separate from the drive unit, space saving, applications in a wide range of environmental conditions.

LABORTECHNIK FRÖBEL

BERLIN - LINDAU
Alwindstraße 4 D-88131 LINDAU
Fax 8382/23120 > 1.1.94: 8382/985232
in Austria - Ing. Neugebauer Vienna
in Switzerland - LAS-Schönengrund

Reader Service No.4



NEW FACILITY ANNOUNCEMENT

Computer Cell Culture Center has the opportunity to build a flexible multiuse facility to grow (recombinant) mammalian cells and to produce (recombinant) viruses on a 8-32L scale using one out of four different suites, each meeting FDA and ISO 9000 specifications. Full biological containment, scientific equipment for purification and analysis will be provided in each suite. Favourable scientific environment. If you are interested in such a facility (for rental or for subcontract), please write or fax to:

A.O.A. Miller
Computer Cell Culture Center
681 Parc de la Sablonnière
B-7000 Mons
Tel: +32 65 37 35 51 Fax: +32 65 36 22 22

4C YOUR PARTNER
IN MASS CULTIVATION OF MAMMALIAN CELLS

Reader Service No.5